medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Full title: Defining Suicidal Thought and Behavior Phenotypes for Genetic Studies

Authors: Eric T. Monson [1,2], Sarah M. C. Colbert [3,4,5], Ole A. Andreassen [6,7], Olatunde O. Ayinde [8], Cosmin A. Bejan [9], Zuriel Ceja [10,11], Hilary Coon [1,2], Emily DiBlasi [1,2], Anastasia Izotova [12,13,14], Erin A. Kaufman [1,2], Maria Koromina [3,4,5], Woojae Myung [15,16], John I. Nurnberger Jr. [17,18], Alessandro Serretti [19], Jordan W. Smoller [20,21,22], Murray B. Stein [23], Clement C. Zai [21,24,25,26,27] for the Suicide Working Group of the Psychiatric Genomics Consortium, Mihaela Aslan [28,29], Peter B. Barr [30,31,32,33], Tim B. Bigdeli [30,31,32,33], Philip D. Harvey [34,35], Nathan A. Kimbrel [36,37,38,39], Pujan R. Patel [36,37] for the Cooperative Studies Program (CSP) #572, Douglas Ruderfer [9,40,41], Anna R. Docherty [1,2,42], Niamh Mullins [3,4,5], J. John Mann [43]

Affiliations:

[1] Department of Psychiatry, University of Utah Spencer Fox Eccles School of Medicine
[2] Huntsman Mental Health Institute
[3] Department of Psychiatry, Icahn School of Medicine at Mount Sinai
[4] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai
[5] Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai
[6] Division of Mental Health and Addiction, Oslo University Hospital
[7] NORMENT Centre, University of Oslo
[8] Department of Psychiatry, College of Medicine, University of Ibadan
[9] Department of Biomedical Informatics, Vanderbilt University Medical Center
[10] Mental Health and Neuroscience Program, QIMR Berghofer Medical Research Institute
[11] School of Biomedical Sciences, Faculty of Medicine, The University of Queensland
[12] Nic Waals Institute, Lovisenberg Diaconal Hospital
[13] PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health
[14] Department of Psychology, University of Oslo
[15] Department of Neuropsychiatry, Seoul National University Bundang Hospital
[16] Department of Psychiatry, Seoul National University College of Medicine
[17] Department of Psychiatry, Indiana University School of Medicine
[18] Department of Medical & Molecular Genetics, Indiana University
[19] Department of Medicine and Surgery, Kore University of Enna, Italy
[20] Center for Precision Psychiatry, Department of Psychiatry, Massachusetts General Hospital
[21] Stanley Center for Psychiatric Research, Broad Institute
[22] Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Massachusetts General Hospital
[23] Department of Psychiatry and School of Public Health, University of California San Diego
[24] Department of Psychiatry, University of Toronto
[25] Institute of Medical Science, University of Toronto
[26] Molecular Brain Science, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health
[27] Laboratory Medicine and Pathobiology, University of Toronto

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[28] Cooperative Studies Program Clinical Epidemiology Research Center (CSP-CERC), VA Connecticut Healthcare System
[29] Department of Internal Medicine, Yale University School of Medicine
[30] Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University
[31] VA New York Harbor Healthcare System
[32] Institute for Genomics in Health, SUNY Downstate Health Sciences University
[33] Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University
[34] Bruce W. Carter Miami Veterans Affairs (VA) Medical Center
[35] University of Miami School of Medicine
[36] Durham VA Health Care System
[37] VA Health Services Research and Development Center of Innovation to Accelerate Discovery and Practice Transformation
[38] VISN 6 Mid-Atlantic Mental Illness Research, Education, and Clinical Center
[39] Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine
[40] Vanderbilt Genetics Institute, Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center
[41] Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center
[42] Clinical and Translational Science Institute & the Center for Genomic Medicine, University of Utah
[43] Departments of Psychiatry and Radiology, Columbia University

Corresponding Author Contact: Eric T. Monson, eric.monson@hsc.utah.edu

Running title: Defining Suicidal Thought and Behavior Phenotypes

Keywords: Suicidality, suicide attempt, suicidal ideation, phenotype, International Classification of Disease, Psychiatric Genetics

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background:

Standardized definitions of suicidality phenotypes, including suicidal ideation (SI), attempt (SA), and death (SD) are a critical step towards improving understanding and comparison of results in suicide research. The complexity of suicidality contributes to heterogeneity in phenotype definitions, impeding evaluation of clinical and genetic risk factors across studies and efforts to combine samples within consortia. Here, we present expert and data-supported recommendations for defining suicidality and control phenotypes to facilitate merging current/legacy samples with definition variability and aid future sample creation.

Methods:

A subgroup of clinician researchers and experts from the Suicide Workgroup of the Psychiatric Genomics Consortium (PGC) reviewed existing PGC definitions for SI, SA, SD, and control groups and generated preliminary consensus guidelines for instrument-derived and international classification of disease (ICD) data. ICD lists were validated in two independent datasets (N = 9,151 and 12,394).

Results:

Recommendations are provided for evaluated instruments for SA and SI, emphasizing selection of lifetime measures phenotype-specific wording. Recommendations are also provided for defining SI and SD from ICD data. As the SA ICD definition is complex, SA code list recommendations were validated against instrument results with sensitivity (range = 15.4% to 80.6%), specificity (range = 67.6% to 97.4%), and positive predictive values (range = 0.59-0.93) reported.

Conclusions:

Best-practice guidelines are presented for the use of existing information to define SI/SA/SD in consortia research. These proposed definitions are expected to facilitate more

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

homogeneous data aggregation for genetic and multisite studies. Future research should involve refinement, improved generalizability, and validation in diverse populations.

Introduction

The complex nature of suicidality and non-suicidal self-harm phenotypes has contributed to diverse definitions that hamper comparison and reproducibility across studies. Consistent suicidality definitions would allow for more robust and generalizable studies. Specifically, suicidal ideation (SI), suicide attempt (SA) and death by suicide (SD) represent complex and partially overlapping phenotypes, collectively referred to as suicidality. Importantly, suicidality is distinct from non-suicidal self-injury (NSSI). Table 1 displays current international standard phenotype definitions along with aggregated phenotype names (1,2).

Phenotypic complexity arises from shared and independent risk factors from genetic and environmental sources. SD and SA have an estimated heritability of 30-50% based on family and twin studies (3), with a portion of the heritability arising independently from that of related psychiatric disorders (4). SI is also heritable (estimated at 36-43% from twin and family studies) (5), but has less estimated independent heritability from psychiatric disorders (6). Results of genome-wide association studies (GWAS) have found genetic correlation ($r_g \pm$ standard error) between suicidality phenotypes and NSSI, but estimates of overlap have varied. For example, between SA and SD (r_g = 0.69±0.15) (7), SI and SA (r_g = 0.71±0.09) (8), and SA and NSSI (r_g = 0.59±0.11 (9), r_g = 0.99±0.16 (8)). A recent study of US Army soldiers found that polygenic risk score (PRS) for SA was predictive of lifetime SA but was not predictive of lifetime NSSI (10). Like most psychiatric disorders, much of the genetic liability for each suicidality phenotype arises from numerous common and rare genetic variants. The polygenic architecture means that hundreds of thousands of samples are required to conduct well-powered GWAS, necessitating the formation of consortia, such as the Psychiatric Genomics Consortium (PGC), to conduct meta-analyses across cohorts (4,11) . These consortia datasets also encompass a

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

wide range of ancestral diversity, facilitating the generalizability of results across global populations.

Importantly, consortia combine numerous legacy cohorts with considerable variability in phenotype definitions. Reduction of heterogeneity from varied ascertainment allows for more powerful and productive comparisons. Implementing consistent phenotype definitions across cohorts could considerably reduce heterogeneity. Optimally, phenotype definitions should be focused on ease of implementation, to aid in consistency and adoption, but also be flexible enough to allow varied data types, including instrument and electronic health record (EHR) data. Such definitions would substantially benefit collaborative efforts by increased sample sizes, comparability, statistical power, reproducibility, and opportunities for meta-analysis across studies.

Flexible phenotypes may also help overcome limitations of current diagnostic definitions and coding options. For example, a proposed diagnosis of “Suicidal Behavior Disorder” has been added within the section of the Diagnostic and Statistical Manual of Mental Health, revision 5 (DSM-5) (12) entitled “conditions for further study”, but is not yet formally defined. International Classification of Disease (ICD) (13) descriptives (such as V codes in ICD 9 and R and Z codes in ICD 10) e.g. Z91.51 "personal history of suicidal behavior" can be very helpful, but are not frequently used (14,15). In the meantime, aligning phenotypes with ICD code lists that providers utilize may allow this gap to be bridged and facilitate international research.

To address these challenges, we propose a flexible set of guidelines to represent best practices in defining suicidality and control phenotypes to be adopted within the Psychiatric Genomics Consortium Suicide Working Group and which can be implemented within the field. Specifically, this protocol provides recommendations for utilization of instrument, public health, and EHR data. While strategies for employing this protocol must take into account the primary study questions and goals, the utilization of these recommendations will allow for more

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

consistent study design throughout the field to improve comparability, power, and reproducibility of results.

Methods and Materials

Considerations for creating a definition of suicidality categories from instruments

Many instruments have been developed for identifying prior SI and SA. These instruments are typically designed for specific target populations, environments. Instruments may also require evaluator training and have variable quality and validation metrics available. Many instruments are meant to be used for triage and prediction of future events, rather than research, and vary widely in performance for prediction (16–20).Therefore, selecting an appropriate instrument for a research study or a clinical setting can be difficult, and involves several considerations.

First, the intended population and purpose of the instrument. Some instruments have been explicitly developed to complement research or forensic efforts. Such measures often consist of detailed interviews and require specialized training to administer to meet validity criteria. These instruments also provide varied, and sometimes limited, information for SI or SA. Examples include the “Composite International Diagnostic Interview” (CIDI)(21), “Diagnostic Interview for Genetic Studies” (DIGS)(22), and the “Structured Clinical Interview for DSM-5” (SCID-5)(23), among many others.

At the other end of the spectrum, many instruments have been designed for rapid, high-sensitivity screening in acute or routine clinical settings. Such measures often ask few and/or time-limited questions to evaluate acute suicide risk for triage purposes. Examples include the “Ask Suicide-Screening Questions” (ASQ) (24) and the “Suicidal Behavior Questionnaire - Revised” (SBQ-R) (25). Other rapid screeners seek to capture a broader mental health snapshot, such as in the widely used “Patient Health Questionnaire” (PHQ) measures (26). Overall, rapid screeners are not typically designed with research efforts in mind, and may obtain

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

a “minimal” phenotype. For example, item 9 in the popular PHQ-9 does not separate SI from non-suicidal self-injury ideation, making it less specific than more detailed evaluations (27,28).

Finally, consideration should be given to the time frame assessed. Instruments may evaluate lifetime or time-limited events and may consider other factors, including frequency, severity, and intent. The “Columbia Suicide Severity Rating Scale” (C-SSRS) (29), developed by the Columbia University group (including J.J.M), represents a widely adopted, fairly comprehensive evaluation. Like many instruments, the C-SSRS has multiple versions that may assess lifetime or limited time windows. Careful evaluation of an instrument for questions such as “in the past month…”, or “during your worst episode of depression” should be performed to be certain of the timeframe assessed and this should be accurately reported. A positive result from a time-limited instrument is suitable for inclusion but a negative result is not always suitable for exclusion. Ideally, all available sources of information on the study population should be leveraged to define phenotypes as precisely as possible.

Considerations for Creating a definition of suicidality categories from ICD data

Among suicide phenotypes, SD and SI are the most straightforward to define using ICD codes. Specifically, SD is defined using cause of death (COD) ICD codes assigned by a coroner or medical examiner based on available information. Evidence indicates that factors such as globally varying autopsy rates, inconsistent patterns of use of ICD codes by medical examiners or coroners, and uncertainty about the intent surrounding the death, may contribute to “missed” cases (30–32). However, in the absence of additional resources such as death registry data, psychological autopsy, or family interview data, it is recommended that researchers classify SD strictly according to the COD codes. For SI, only a single code is defined in both ICD-9 (V62.84) and ICD-10 (R45.851). These SI codes may be infrequently used, leading to missed events (33). However, SI codes are the only option for defining SI when other data sources, such as EHR data, are unavailable.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

For SA, however, there are many ICD codes available for consideration, and there are many challenges to utilizing these. First, ICD codes are considerably less sensitive than instruments or evolving methods based on natural language processing of narrative clinical free text notes (34–36) , where specific events may be mentioned that are not coded. Other ICD complexities include: 1) lack of a uniform standard and scope of practice for clinician coding training; 2) variable clarity, accuracy, suicidal intent, and completeness of patient history; 3) time limitations of the provider; 4) insurance billing requirements and provider policies that may influence the use of some codes for administrative purposes; 5) differences in region, population, religion, and culture that influence stigma regarding SA; and 6) potential liability or legal implications of assigning an SA code (37). Despite these limitations, ICD codes represent a conservative and important source of data in large and diverse public health and population cohorts worldwide and may be the only data source available.

As noted above, an effort to develop a single list of ICD codes for SA was published by the National Center for Health Statistics (NCHS) (38). However, some codes on this list may better represent NSSI. In addition, all codes that contain the qualifier “undetermined intent” or “accidental” for a reported injury were omitted from the NCHS SA list, although a proportion of these will capture SA.

The NCHS list was used as a template, along with the considerations above, for generating ICD guideline lists. Two lists were generated, representing 1) an SA definition and 2) an SA screening list that includes NSSI and codes of uncertain intent. ICD-10 and ICD-10-CM (United States) code lists were obtained from https://icd.who.int/browse10/2019/en#/ and https://www.cdc.gov/nchs/icd/Comprehensive-Listing-of-ICD-10-CM-Files.htm, respectively. ICD-9 and ICD-9-CM code lists were obtained from https://apps.who.int/iris/handle/10665/39473 and https://www.cdc.gov/nchs/icd/icd9cm.htm , respectively.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Suicide Attempt ICD List Validation

The codes in the SA definition and screening lists were validated in two datasets with ICD and C-SSRS, or equivalent instrument data, available. The NCHS published ICD code list (38) was also evaluated for comparison in these data. For each of the three lists, specificity, sensitivity, and positive predictive value (PPV) were calculated for the detection of prior SA as compared with instrument responses. In addition, results were stratified by sex.

Both of the validation cohorts have been previously described, in detail. Briefly, the "Cooperative Studies Program" (CSP) #572 cohort, entitled "Genetics of Functional Disability in Schizophrenia and Bipolar Illness Validation" was collected to evaluate genetic and other characteristics of veterans with severe psychiatric illness (39,40). Participants in CSP #572 (N = 9,151, 86.2% male) universally received C-SSRS screening during the baseline diagnostic interview. The Vanderbilt University Medical Center (VUMC) cohort was obtained in one of two ways. First, through de-identified VUMC medical record data from a research-oriented data repository, Synthetic Derivative (41). Synthetic Derivative cases that appeared to be enriched for suicidal behavior were selected for manual validation through review of electronic health record clinical notes (N = 1,095, 43.2% male, 56.8% female) (35). Second, via deidentified individuals seen at a VUMC psychiatric assessment service, all of whom received suicide risk screening (N = 11,299, 47.6% male, 52.4% female) (35). All data were handled in accordance with oversight and given ethical approval from the Vanderbilt University Medical Center and Veterans Affairs Central Institutional Review Boards, as noted in the respective referenced works. It is also noted that all data were processed within responsible groups with only the presented summary data being shared.

Considerations for creating a definition for control samples

Appropriate selection of a control sample for studies examining suicide phenotypes should consider: 1) primary study question(s); 2) data availability, such as ICD codes or

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

instrument data; 3) sample size; and 4) ascertainment strategies and variables within control groups, such as population background demographics, and treatment setting. A specific issue that requires consideration in suicidality studies is the strong underlying association with other psychiatric diagnoses, present in an estimated 60-98% of cases (42). SA and SD have a low base rate in the general population, with an approximate annual incidence rate of 0.5% for SA and 0.014% for SD within the United States in 2021 (43), and worldwide SA lifetime prevalence of 2.7% (44). Psychiatric diagnoses have an estimated worldwide general population prevalence of ~30% (45) and 90% within individuals with a history of SA (46). Therefore, controls with psychiatric diagnoses may assist in estimating the effects of these common diagnoses (Figure 1A). However, screening specifically for SI, SA, SD, and NSSI may serve to improve the power of comparisons without introducing significant bias (Figure 1B) (47).

Other factors, such as ascertainment and intensity of screening, must also be considered. Control sample ascertainment based on voluntary response can lead to underestimation of adverse health outcomes and skewed demographic makeup (48). Similarly, intensive screening, such as screening for all psychiatric disorders, can lead to bias (Figure 1C). Such biases impede the parsing of results to isolate the effect of a given suicidality phenotype versus psychiatric illnesses and other comorbidities (49). Such extreme comparisons may also limit generalizability through unintended enrichment of demographic or other sample differences, potentially highlighting spurious or clinically irrelevant associations (50). Finally, in cases where explained variance within a population is a key measure, as in many large-scale genetic studies, overly-screened controls may distort this estimate (49) and also cause distortions in genetic correlation estimates (51).

Fortunately, there are strategies to assess the impact of potentially biased control groups, as detailed in Figure 2. In scenarios where researchers have multiple control samples, performing multiple rounds of analysis may allow evaluation of the impact of any bias. For example, controls least suspected of bias are analyzed, followed by adding samples suspected

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

to have greater bias/less reliable definitions, or performing leave-one-out analyses. Alternatively, strategies such as inverse probability weighting can be utilized to address potential bias effects by weighting observations on the probability of selection to a given comparison group (52). Implemented strategy and all results should be clearly described in published materials.

Results

Questionnaire suicidality definitions

A total of 50 instruments, including unique versions, were assessed during the development of this protocol. These were specifically selected for review due to their use in one or more cohorts included in the PGC, and are not an exhaustive list of possible instruments. Items evaluated for defining SA and SI are presented within supplemental Tables S1 and S2, respectively. Specific factors were considered in evaluation, including clarity of language, whether single or multiple phenotypes were assessed in a single item, and time interval assessed. It is noted that several instruments have multiple versions, some of which are time-limited, and researchers should identify the timeframe assessed by any instrument used. Additionally, frequently used instruments that define a minimal/less-preferred phenotype were also considered.

Based on expert consensus, it is recommended that new studies strongly consider an instrument focused on a detailed assessment of suicidality (SI/SA/SD), wide distribution and language availability, and one with broadly utilized definitions. The Columbia Suicide Severity Rating Scale (C-SSRS) meets these criterion, and it is recommended that this or a similarly constructed instrument be employed. In addition, it is recommended that versions of instruments that assess lifetime history of suicidality be utilized, where possible.

ICD Suicide Attempt definition

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

A classification system was developed by a team of international expert psychiatrists (Mann, Monson, Serretti, Smoller, Sokolowski, Stein) who reviewed the ICD codes with consideration of sensitivity, specificity, potential for misclassification and acceptable error rate. Given the complexity of ascribing suicidal intent to any given ICD code, two lists of classification were generated: the SA definition and screening lists, designed to represent more specific versus more sensitive models, respectively.

The generated ICD lists are summarized in Tables 2 and 3. Complete lists of ICD codes are available in Tables S3 and S4. It is recommended that a single occurrence of any of the listed codes in a patient's EHR be considered a positive result for SA. This is recommended on the basis of low prevalence of SA combined with infrequent SA coding within the clinical setting.

Consideration was given regarding the likely intent of different coded events. Specifically, deliberate self-harm injury that presents for medical evaluation may represent SA or NSSI. NSSI is common, with prevalence estimates of 4% in the general adult population, 20% in adult psychiatric patient populations (53), 13-17% in adolescents and young adults (54). NSSI is a risk factor for SA and SD, with possible increased risk in women (55) and in younger populations, with decreasing NSSI frequency as age increases (56).

Phenomenological characteristics of NSSI can be used to separate it from SA. For example, NSSI typically involves repetitive cutaneous injuries like cutting, burning, or banging the head or part of the body against something hard (57,58). Specifically, arm/wrist cutting is less likely to predict future SA/SD (59) and open superficial wounds of the forearm are rarely associated with SD but frequently with NSSI (60). Stabbing and serious cutting injuries are also much rarer than superficial cutaneous injuries (61). Therefore, it was decided that exclusion of codes specific to superficial cutaneous cutting and burning injuries would be an acceptable tradeoff of sensitivity and specificity for the definition list.

Other codes more frequently associated with suicidality than with NSSI include poisoning via carbon monoxide and hormonal, antiepileptic, pain, and psychotropic medication

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

overdose, as well as other methods like jumping from a height or in front of a train or vehicle, hanging, self-inflicted gunshot wound (61). In addition, insulin overdose cases have been overwhelmingly observed to be related to SA (rates of 89-95%) (62,63). Overdosing on antidepressants, especially in combination with controlled substances, also correlates with SD (64). Drowning deaths may also be subject to possible misclassification in intent throughout the world (65). Suicidal intent in prisoners is correlated with the seriousness of injury; hanging was correlated with SA in 80% of the cases while cutting or striking one's head rarely constituted SA (19% for superficial cuts, 21% for all cuts, and 0% for head-striking) (66). ICD codes were added to the definition and/or screening lists for SA considering these tendencies, noting sensitivity gains and limited specificity tradeoff (67).

ICD Suicide Attempt Definition Validation

Validation of SA lists in the CSP #572 and VUMC samples are provided in Table 4. Note that two time points were provided with the CSP #572 sample: ICD codes assigned pre and post study entrance. Including multiple CSP #572 time points and VUMC clinical settings allows evaluation of how sensitivity and specificity of ICD codes may change based on clinical setting and emphasis. Across all settings, the developed SA definition ICD list had similar or greater PPV than the existing NCHS list (Table 4). The SA screening list demonstrated consistently higher sensitivity compared with other definition lists, but had the lowest PPV (Table 4). Notably, the sensitivity of ICD codes varied considerably based on population/clinical setting and the list used, ranging from 15.20% to 71.09% of subjects captured by instruments. Sex-stratified secondary analyses in CSP #572 and VUMC found that in general, ICD sensitivity, PPV, and specificity values were similar between males and females within both samples (Table S5).

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Discussion

SI, SA, and SD are complex phenotypes that are partially overlapping, and also related to, but distinct from, NSSI. These phenotypes primarily differ in terms of intent, a factor that is often not trivial for providers to determine. Many studies utilize varied and/or aggregate phenotypes (e.g., “suicidality”) to study these phenomena. Together these factors have led to considerable heterogeneity in research phenotype definitions, which poses challenges to comparability, replication, and meta-analyses across studies.

Our proposed guidelines attempt to address this issue by providing best-practice guidance for defining SI, SA and SD using common data sources, including ICD codes, instruments, and public health information. As observed in the validation of ICD codes, instrument responses, used as the baseline for the validation, had much greater sensitivity than ICD codes for SA, as anticipated. Specifically, codes captured, at best, 70% of cases and frequently less than 50% in the populations assessed that instruments would have captured. Therefore, the use of an appropriate instrument, as outlined here, is strongly recommended in the design of future studies to investigate SA.

Instruments that adequately assess suicidality across the lifetime and which have been produced and validated in the native language of assessed individuals are strongly preferred. However, in some cases, rapid screening measures may be useful, as listed in the “acceptable/minimal” segment of Tables S1 and S2. Such inclusion of “minimal phenotypes” must be balanced against a potential loss of specificity and risk of drawing conclusions that may only apply to aggregated phenotypes, as has been described in major depression (68). Of note, the impact of the inclusion of less specific phenotypes or biased controls can and should be assessed in many ways, such as leave-one-out or inverse-weighted meta-analyses, examination of heritability, genetic correlation, and heterogeneity of effect sizes (69,70).

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

If instrument or EHR data are not available, the presented ICD lists give additional options for assessing data.The SA definition list had high PPV (0.71 - 0.92) and the SA screening list had higher sensitivity (17.0% - 71.1%). Therefore, the SA definition could be used as a primary phenotype and the SA screening as a method to screen the control sample. Alternatively, the SA screening tier could be used to generate a more sensitive phenotype to maximize sample size for a study, though with more error. The variation in performance between the validation cohorts likely reflects variable usage of ICD codes in differing clinical settings and populations, which can be hard to predict, but should be considered when interpreting results.

Ultimately, the optimal strategy to define a given suicide phenotype may vary based on available information, sample size, and the study question. It is important to reiterate that instrument responses are much more sensitive than ICD codes for the screening of suicide phenotypes, and are of the highest value to include if constructing a new cohort. In addition, including consideration of other data elements, if available, including method and/or lethality of attempt, duration of suicidality, and age at first attempt may allow for even more strict definitions and evaluation of more specific phenotypes. Regardless of the strategy employed, a clear explanation of the rationale and design of the sample will aid future replication and meta-analysis efforts.

Study Limitations

The defined guidelines have potential limitations. These guidelines were designed to be approachable to diverse researchers and are based on existing work. As such, only a selection of available suicidality instruments were evaluated within these guidelines. There will inevitably be data collected using alternative instruments and these should be evaluated using the general principles provided here.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Similarly, alternative strategies that make use of more extensive EHR data, such as natural language processing and manual review, are not evaluated in this work. These methods are not yet standardized and require more extensive access to individual records or resources than many studies may have. Future iterations of this protocol may include consideration of these strategies.

The use of simple list definitions of individual ICD codes may also reduce the capacity to differentiate SA and NSSI features compared with classification models that use multiple codes or EHR elements (71). A simple design was selected intentionally to increase portability and usability.

Finally, the ICD code validation was also performed only within United States samples due to limitations in existing non-US samples with the required data types, limiting generalization in international samples. Even other US samples may vary considerably due to variations in clinical documentation practices and systemic biases, including racial bias, in psychiatric diagnoses (72).

Conclusion

These guidelines have been designed to improve the overall consistency of phenotyping and sample selection for ongoing and future suicide research studies. They are intended to serve as a framework for the design of genetic and consortium suicidality studies, and will not be ideal for every study. However, the evaluation and discussion of the complexity of suicide-related phenotypes is a crucial consideration when establishing future samples and criteria. In addition, it is strongly recommended that any effort exploring a suicide phenotype will provide clear descriptions of selection criteria and rationale for the design on the basis of the study question with these complexities in mind. Improving the clarity and consistency of samples will be critical to identify robust and interpretable risk factors for SI/SA/SD and related phenotypes,

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

in an effort to improve the identification of high-risk individuals, improve treatment modalities, reduce associated tangible and personal costs, and ultimately prevent these outcomes.

Acknowledgements

This work was funded by NIMH R01MH124839 (Mullins), NIMH R01MH132733 (Ruderfer), R01MH123619 (Docherty), R01MH123489 (Coon), Brain and Behavior Research Foundation (NARSAD Young investigator award No. 31248 to E.M.), and the Huntsman Mental Health Institute. This work was also supported by significant funding that has been provided to the Psychiatric Genomics Consortium from the National Institute of Mental Health (PGC3: U01 MH109528; PGC2: U01 MH094421; PGC1: U01 MH085520). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1842169. JWS was supported in part by R01 MH118233 and MH117599.

Additional Members of the Suicide Working Group of the Psychiatric Genomics Consortium: Mark Adams, Rolf Adolfsson, Ingrid Agartz, Esben Agerbo, Tracy M Air, Martin Alda, Lars Alfredsson, Adebayo Anjorin, Vivek Appadurai, María Soler Artigas, Allison E. Ashley-Koch, Swapnil Awasthi, M Helena Azevedo, Amanda Bakian, Nicholas Bass, Claiton HD Bau, Bernhard T Baune, Jean C. Beckham, Frank Bellivier, Andrew W Bergen, Klaus Berger, Wade H Berrettini, Joanna M Biernacka, Elisabeth B Binder, Michael Boehnke, Martin Bohus, Marco P Boks, Anders D Børglum, Rosa Bosch, David L Braff, Harry Brandt, Gerome Breen, Richard Bryant, Monika Budde, Cynthia M Bulik, Enda M Byrne, Wiepke Cahn, Adrian I Campos, Miguel Casas, Enrique Castelao, Jorge A Cervilla, Xiao Chang, Boris Chaumette, Hsi-Chung Chen, Wei J Chen, Erik D Christensen, Sven Cichon, Jonathan R I Coleman, Aiden Corvin, Nicholas Craddock, David Craig, Steven Crawford, Scott Crow, Franziska Degenhardt, Ditte Demontis, Michelle Dennis, Srdjan Djurovic, Philibert Duriez, Howard J Edenberg, Alexis C Edwards, Annette Erlangsen, Tõnu Esko, Ayman H Fanous, Fernando Fernández-Aranda, Manfred M

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Fichter, Jerome C Foo, Andreas J Forstner, Mark Frye, Janice M Fullerton, Hanga Galfalvy, Steven Gallinger, Michael Gandal, Melanie Garrett, Justine M Gatt, Pablo V Gejman, Joel Gelernter, Ina Giegling, Stephen J Glatt, Philip Gorwood, Hans J Grabe, Melissa J Green, Eugenio H Grevet, Maria Grigoroiu-Serbanescu, Yiran Guo, Blanca Gutierrez, Jose Guzman-Parra, Jonathan D Hafferty, Lauren Hair, Hakon Hakonarson, Katherine A Halmi, Steven P Hamilton, Marian L Hamshere, Annette M Hartmann, Elizabeth R. Hauser, Michael A. Hauser, Joanna Hauser, Stefanie Heilmann-Heimbach, Akitoyo Hishimoto, Per Hoffmann, David M Hougaard, Jennifer Huffman, Hai-Gwo Hwu, Marcus Ising, Daniel Jacobson, Sonia Jain, Stéphane Jamain, Susana Jiménez-Murcia, Craig Johnson, Ian Jones, Lisa A Jones, Lina Jonsson, René S Kahn, JooEun Kang, Allan S Kaplan, Walter H Kaye, Pamela K Keel, John R Kelsoe, Kenneth S Kendler, James L Kennedy, Ronald C Kessler, Minsoo Kim, Stefan Kloiber, Kelly L Klump, Karestan C Koenen, Manolis Kogevinas, Bettina Konte, Henry R Kranzler, Marie-Odile Krebs, Po-Hsiu Kuo, Mikael Landén, Jacob Lawrence, Marion Leboyer, Phil H Lee, Daniel F Levey, Douglas F Levinson, Cathryn M Lewis, Dong Li, Qingqin S Li, Shih-Cheng Liao, Calwing Liao, Klaus Lieb, Lisa Lilenfeld, Jennifer H. Lindquist, Jolanta Lissowska, Chih-Min Liu, Adriana Lori, Susanne Lucae, Ravi Madduri, Pierre J Magistretti, Christian R Marshall, Nicholas G Martin, Fermin Mayoral, Susan L McElroy, Patrick McGrath, Peter McGuffin, Andrew M McIntosh, Benjamin McMahon, Andrew McQuillin, Sarah E Medland, Divya Mehta, Ingrid Melle, Yuri Milaneschi, James E Mitchell, Philip B Mitchell, Esther Molina, Gunnar Morken, Ole Mors, Preben Bo Mortensen, Bertram Müller-Myhsok, Richard M Myers, Caroline Nievergelt, Vishwajit Nimgaonkar, Merete Nordentoft, Markus M Nöthen, Michael C O'Donovan, Satoshi Okazaki, Catherine M Olsen, Roel A Ophoff, David W. Oslin, Ikuo Otsuka, Michael J Owen, Carlos Pato, Michele T Pato, Brenda WJH Penninx, Jonathan Pimm, Dalila Pinto, Giorgio Pistis, Renato Polimonti, David Porteous, James B Potash, Robert A Power, Abigail Powers, Martin Preisig, Xuejun Qin, Digby Quested, Josep Antoni Ramos-Quiroga, Nicolas Ramoz, Andreas Reif, Miguel E Rentería, Marta Ribasés, Vanesa Richarte, Marcella Rietschel, Stephan Ripke,

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Margarita Rivera, Andrea Roberts, Gloria Roberts, Stefan Roepke, Guy A Rouleau, Diego L Rovaris, Vsevolod Rozanov, Dan Rujescu, Cristina Sánchez-Mora, Alan R Sanders, Stephen W Scherer, Christian Schmahl, Peter R Schofield, Thomas G Schulze, Laura J Scott, Andrey Shabalin, Jianxin Shi, Stanley I Shyn, Lea Sirignano, Pamela Sklar, Olav B Smeland, Daniel J Smith, Marcus Sokolowski, Edmund J S Sonuga-Barke, Gianfranco Spalletta, Eli A Stahl, Anna Starnawska, John S Strauss, Fabian Streit, Michael Strober, Mei-Hsin Su, Beata Świątkowska, Laura M Thornton, Jodie Trafton, Janet Treasure, Maciej Trzaskowski, Ming T Tsuang, Gustavo Turecki, Robert J Ursano, Sandra Van der Auwera, Laura Vilar-Ribó, John B Vincent, Henry Völzke, Consuelo Walss-Bass, James TR Walters, Erin B Ware, Danuta Wasserman, Hunna J Watson, Cynthia Shannon Weickert, Thomas W Weickert, Myrna M Weissman, Frank Wendt, Thomas Werge, David C Whiteman, Leanne M Williams, Virginia Willour, Stephanie H Witt, D Blake Woodside, Naomi R Wray, Zeynep Yilmaz, Lea Zillich.

The CSP #572 study team

Planning Committee: M. Aslan, M. Antonelli, M. de Asis, M. S. Bauer, M. Brophy, J. Concato, F. Cunningham, R. Freedman, M. Gaziano, T. Gleason, P. D. Harvey, G. Huang, J. Kelsoe, T. Kosten, T. Lehner, J. B. Lohr, S. R. Marder, P. Miller, T.J. O'Leary, T. Patterson, P. Peduzzi, R. Przygodzki, L. Siever, P. Sklar, S. Strakowski, H. Zhao.

Executive Committee: M. Brophy, J. Concato, A.H. Fanous, M. Gaziano, P.D. Harvey, T. Kosten, A. Malhotra, S. Mane, T. Bigdeli, P. Sklar, L. Siever, H. Zhao.

Study Chairs' Offices: VA Healthcare System, Bronx, NY: L. Siever (Co-Chair), M. Corsey, L. Zaluda. VA Healthcare System, Miami, FL: P.D. Harvey (Co-Chair), J. Johnson.

CSP Epidemiology Centers: VA Clinical Epidemiology Research Center CERC, VA Connecticut Healthcare System, West Haven, CT, included J. Concato Director, Methodological Co-Principal Proponent, M. Aslan, D. Cavaliere, V. Jeanpaul, A. Maffucci, L. Mancini; the Massachusetts Veterans Epidemiology Research and Information Center MAVERIC, VA Boston Healthcare

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

System, Jamaica Plain, MA, included M. Gaziano Director, Methodological Co-Principal Proponent, J. Deen, G. Muldoon, S. Whitbourne.

Study Sites: Albuquerque: J. Canive, L. Adamson, L. Calais, G. Fuldauer, R. Kushner, G. Toney, M. Lackey, A. Mank, N. Mahdavi, G. Villarreal. Atlanta: E. C. Muly, F. Amin, M. Dent, J. Wold. Baltimore: B. Fischer, A. Elliott, C. Felix, G. Gill. Birmingham: P. E. Parker, C. Logan, J. McAlpine. Boston/Brockton: L.E. DeLisi,. Charleston: M.B. Hammer, D. Agbor-Tabie, W. Goodson. Cincinnati: M. Aslam, M. Grainger, Neil Richtand, Alexander Rybalsky. Houston: R. Al Jurdi, E. Boeckman, T. Natividad, D. Smith, M. Stewart, S. Torres, Z. Zhao. Indianapolis: A. Mayeda, A. Green, J. Hofstetter, S. Ngombu, M. K. Scott, A. Strasburger, J. Sumner. Little Rock: G. Paschall, J. Mucciarelli, R. Owen, S. Theus, D. Tompkins. Long Beach: S.G. Potkin, C. Reist, M. Novin, S. Khalaghizadeh. Miami: R. Douyon, J. Johnson, N. Kumar, B. Martinez. Minneapolis: S.R. Sponheim, T.L. Bender, H.L. Lucas, A.M. Lyon, M.P. Marggraf, L.H. Sorensen, C.R. Surerus. Montrose: C. Sison, J. Amato, D.R. Johnson, N. Pagan-Howard. New York Harbor: L.A. Adler, S. Alperin, T. Leon. Northampton: K.M. Mattocks, N. Araeva, J.C. Sullivan. Palo Alto: T. Suppes, K. Bratcher, L. Drag, E.G. Fischer, L. Fujitani, S. Gill, D. Grimm, J. Hoblyn, T. Nguyen, E. Nikolaev, L. Shere, R. Relova, A. Vicencio, M. Yip. Philadelphia: I. Hurford, S. Acheampong, G. Carfagno. Pittsburgh: G.L. Haas, C. Appelt, E. Brown, B. Chakraborty, E. Kelly, G. Klima, S. Steinhauer. Salisbury: R.A. Hurley, R. Belle, D. Eknoyan, K. Johnson, J. Lamotte. San Diego: E. Granholm, K. Bradshaw, J. Holden, R. H. Jones, T. Le, I.G. Molina, M. Peyton, I. Ruiz, L. Sally. Tacoma: A. Tapp, S. Devroy, V. Jain, N. Kilzieh, L. Maus, K. Miller, H. Pope, A. Wood. Temple: E. Meyer, P. Givens, P. B. Hicks, S. Justice, K. McNair, J.L. Pena, D.F. Tharp. Tuscaloosa: L. Davis, M. Ban, L. Cheatum, P. Darr, W. Grayson, J. Munford, D. Smith, B. Whitfield, E. Wilson. Washington DC: A.H. Fanous, S.E. Melnikoff, B.L Schwartz, M.A. Tureson. West Haven: D. D'Souza, K. Forselius, M. Ranganathan, L. Rispoli.

Albuquerque, NM, CSP Coordinating Center: M. Sather Director, C. Colling, C. Haakenson, D. Krueger.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

VA Office of Research and Development: T. O'Leary Chief Research and Development Officer, G. Huang Director, Cooperative Studies Program, T. Gleason Director, Clinical Science Research and Development Service, R. Przygodzki Associate Director for Genomic Medicine, and Acting Director of Biomedical Laboratory Research and Development Service, S. Muralidhar Senior Scientific Program Manager Genomic Medicine Program, Biomedical and Clinical R&D Services.

Million Veteran Program: Consortium Acknowledgement for Manuscripts

MVP Executive Committee:

- Co-Chair: J. Michael Gaziano, M.D., M.P.H.
- Co-Chair: Rachel Ramoni, D.M.D., Sc.D.
- Jim Breeling, M.D. ex-officio
- Kyong-Mi Chang, M.D.
- Grant Huang, Ph.D.
- Sumitra Muralidhar, Ph.D.
- Christopher J. O'Donnell, M.D., M.P.H.
- Philip S. Tsao, Ph.D.

MVP Program Office

- Sumitra Muralidhar, Ph.D.
- Jennifer Moser, Ph.D.

MVP Recruitment/Enrollment

- Recruitment/Enrollment Director/Deputy Director, Boston
- Stacey B. Whitbourne, Ph.D.; Jessica V. Brewer, M.P.H.
- MVP Coordinating Centers

– Clinical Epidemiology Research Center CERC, West Haven – John Concato, M.D., M.P.H.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

– Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, Albuquerque

- Stuart Warren, J.D., Pharm D.; Dean P. Argyres, M.S.

– Genomics Coordinating Center, Palo Alto – Philip S. Tsao, Ph.D.

– Massachusetts Veterans Epidemiology Research Information Center MAVERIC, Boston

- J. Michael Gaziano, M.D., M.P.H.

– MVP Information Center, Canandaigua – Brady Stephens, M.S.

- Core Biorepository, Boston – Mary T. Brophy M.D., M.P.H.; Donald E. Humphries, Ph.D.

- MVP Informatics, Boston – Nhan Do, M.D.; Shahpoor Shayan

- Data Operations/Analytics, Boston – Xuan-Mai T. Nguyen, Ph.D.

MVP Science

- Genomics - Christopher J. O'Donnell, M.D., M.P.H.; Saiju Pyarajan Ph.D.; Philip S. Tsao, Ph.D.

- Phenomics - Kelly Cho, M.P.H, Ph.D.

- Data and Computational Sciences – Saiju Pyarajan, Ph.D.

- Statistical Genetics – Elizabeth Hauser, Ph.D.; Yan Sun, Ph.D.; Hongyu Zhao, Ph.D.

MVP Local Site Investigators

- Atlanta VA Medical Center Peter Wilson - Bay Pines VA Healthcare System Rachel McArdle

- Birmingham VA Medical Center Louis Dellitalia

- Cincinnati VA Medical Center John Harley

- Clement J. Zablocki VA Medical Center Jeffrey Whittle

- Durham VA Medical Center Jean Beckham

- Edith Nourse Rogers Memorial Veterans Hospital John Wells

- Edward Hines, Jr. VA Medical Center Salvador Gutierrez

- Fayetteville VA Medical Center Gretchen Gibson

- VA Health Care Upstate New York Laurence Kaminsky

- New Mexico VA Health Care System Gerardo Villareal

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

- VA Boston Healthcare System Scott Kinlay
- VA Western New York Healthcare System Junzhe Xu
- Ralph H. Johnson VA Medical Center Mark Hamner
- Wm. Jennings Bryan Dorn VA Medical Center Kathlyn Sue Haddock
- VA North Texas Health Care System Sujata Bhushan
- Hampton VA Medical Center Pran Iruvanti
- Hunter Holmes McGuire VA Medical Center Michael Godschalk
- Iowa City VA Health Care System Zuhair Ballas
- Jack C. Montgomery VA Medical Center Malcolm Buford
- James A. Haley Veterans' Hospital Stephen Mastorides
- Louisville VA Medical Center Jon Klein
- Manchester VA Medical Center Nora Ratcliffe
- Miami VA Health Care System Hermes Florez
- Michael E. DeBakey VA Medical Center Alan Swann
- Minneapolis VA Health Care System Maureen Murdoch
- N. FL/S. GA Veterans Health System Peruvemba Sriram
- Northport VA Medical Center Shing Shing Yeh
- Overton Brooks VA Medical Center Ronald Washburn
- Philadelphia VA Medical Center Darshana Jhala
- Phoenix VA Health Care System Samuel Aguayo
- Portland VA Medical Center David Cohen
- Providence VA Medical Center Satish Sharma
- Richard Roudebush VA Medical Center John Callaghan
- Salem VA Medical Center Kris Ann Oursler
- San Francisco VA Health Care System Mary Whooley
- South Texas Veterans Health Care System Sunil Ahuja

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

- Southeast Louisiana Veterans Health Care System Amparo Gutierrez
- Southern Arizona VA Health Care System Ronald Schifman
- Sioux Falls VA Health Care System Jennifer Greco
- St. Louis VA Health Care System Michael Rauchman
- Syracuse VA Medical Center Richard Servatius
- VA Eastern Kansas Health Care System Mary Oehlert
- VA Greater Los Angeles Health Care System Agnes Wallbom
- VA Loma Linda Healthcare System Ronald Fernando
- VA Long Beach Healthcare System Timothy Morgan
- VA Maine Healthcare System Todd Stapley
- VA New York Harbor Healthcare System Scott Sherman
- VA Pacific Islands Health Care System Gwenevere Anderson
- VA Palo Alto Health Care System Philip Tsao
- VA Pittsburgh Health Care System Elif Sonel
- VA Puget Sound Health Care System Edward Boyko
- VA Salt Lake City Health Care System Laurence Meyer
- VA San Diego Healthcare System Samir Gupta
- VA Southern Nevada Healthcare System Joseph Fayad
- VA Tennessee Valley Healthcare System Adriana Hung
- Washington DC VA Medical Center Jack Lichy
- W.G. Bill Hefner VA Medical Center Robin Hurley
- White River Junction VA Medical Center Brooks Robey
- William S. Middleton Memorial Veterans Hospital Robert Striker

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Disclosures

Ole Andreassen: Consultant to Cortechs.ai and Precision-Health.ai, and received speaker's honorarium from Lundbeck, Sunovion, Janssen and Otsuka.

Murray Stein: MBS has in the past 3 years received consulting income from Aptinyx, atai Life Sciences, BigHealth, Biogen, Bionomics, Boehringer Ingelheim, Delix Therapeutics, EmpowerPharm, Engrail Therapeutics, Janssen, Jazz Pharmaceuticals, Karuna Therapeutics, Lykos Therapeutics, NeuroTrauma Sciences, Otsuka US, PureTech Health, Sage Therapeutics, Seaport Therapeutics, and Roche/Genentech. Dr. Stein has stock options in Oxeia Biopharmaceuticals and EpiVario. He has been paid for his editorial work on Depression and Anxiety (Editor-in-Chief), Biological Psychiatry (Deputy Editor), and UpToDate (Co-Editor-in-Chief for Psychiatry). He is on the scientific advisory board of the Brain and Behavior Research Foundation and the Anxiety and Depression Association of America.

Philip Harvey: Dr. Harvey consults for a variety of pharmaceutical and device manufacturers on phase 2 and 3 studies of cognition and negative symptoms in SMI. These activities have been reviewed and determined not to be related to the content of this paper.

John Mann: Dr. Mann receives royalties for commercial use of the C-SSR S from the Research Foundation of Mental Hygiene and from Columbia University for the Columbia Pathways App.

All other listed authors declare no conflicts of interest.

References

1. Crosby A, Ortega LV, Melanson C (2011): Self-directed violence surveillance; uniform definitions and recommended data elements. Retrieved from https://stacks.cdc.gov/view/cdc/11997

2. Cipriano A, Cella S, Cotrufo P (2017): Nonsuicidal Self-injury: A Systematic Review. *Front Psychol* 8: 1946.

3. Mann JJ, Arango VA, Avenevoli S, Brent DA, Champagne FA, Clayton P, *et al.* (2009): Candidate endophenotypes for genetic studies of suicidal behavior. *Biol Psychiatry* 65: 556–563.

4. Docherty AR, Mullins N, Ashley-Koch AE, Qin X, Coleman JRI, Shabalin A, *et al.* (2023): GWAS Meta-Analysis of Suicide Attempt: Identification of 12 Genome-Wide Significant Loci

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

and Implication of Genetic Risks for Specific Health Factors. *Am J Psychiatry* 180: 723–738.

5. Brent DA, Mann JJ (2005): Family genetic studies, suicide, and suicidal behavior. *Am J Med Genet C Semin Med Genet* 133C: 13–24.

6. Dutta R, Ball HA, Siribaddana SH, Sumathipala A, Samaraweera S, McGuffin P, Hotopf M (2017): Genetic and other risk factors for suicidal ideation and the relationship with depression. *Psychol Med* 47: 2438–2449.

7. Li QS, Shabalin AA, DiBlasi E, Gopal S, Canuso CM, FinnGen ISGC, *et al.* (2023): Genome-wide association study meta-analysis of suicide death and suicidal behavior. *Mol Psychiatry* 28: 891–900.

8. Kootbodien T, London L, Martin LJ, Defo J, Ramesar R (2023): The shared genetic architecture of suicidal behaviour and psychiatric disorders: A genomic structural equation modelling study. *Front Genet* 14: 1083969.

9. Colbert SMC, Mullins N, Chan G, Meyers JL, Schulman J, Kuperman S, *et al.* (2023): Polygenic Contributions to Suicidal Thoughts and Behaviors in a Sample Ascertained for Alcohol Use Disorders. *Complex Psychiatry* 9: 11–23.

10. Stein MB, Jain S, Papini S, Campbell-Sills L, Choi KW, Martis B, *et al.* (2024): Polygenic risk for suicide attempt is associated with lifetime suicide attempt in US soldiers independent of parental risk. *J Affect Disord* 351: 671–682.

11. Mullins N, Kang J, Campos AI, Coleman JRI, Edwards AC, Galfalvy H, *et al.* (2022): Dissecting the Shared Genetic Architecture of Suicide Attempt, Psychiatric Disorders, and Known Risk Factors. *Biol Psychiatry* 91: 313–327.

12. American Psychiatric Association D (2013): Diagnostic and statistical manual of mental disorders: DSM-5. academia.edu. Retrieved January 26, 2024, from https://www.academia.edu/download/38718268/csl6820_21.pdf

13. World Health Organization (1992): *The ICD-10 Classification of Mental and Behavioural*

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Disorders: Clinical Descriptions and Diagnostic Guidelines. World Health Organization.

14. Alemi F, Avramovic S, Renshaw KD, Kanchi R, Schwartz M (2020): Relative accuracy of social and medical determinants of suicide in electronic health records. *Health Serv Res* 55 Suppl 2: 833–840.
15. Maksut JL, Hodge C, Van CD, Razmi A, Khau MT (2021): Utilization of Z codes for social determinants of health among medicare fee-for-service beneficiaries, 2019. *Office of Minority Health (OMH)*.
16. Chang BP, Tan TM (2015): Suicide screening tools and their association with near-term adverse events in the ED. *Am J Emerg Med* 33: 1680–1683.
17. Mayes TL, Carmody T, Rush AJ, Nandy K, Emslie GJ, Kennard BD, *et al.* (2023): Predicting suicidal events: A comparison of the Concise Health Risk Tracking Self-Report (CHRT-SR) and the Columbia Suicide Severity Rating Scale (C-SSRS). *Psychiatry Res* 326: 115306.
18. Simpson SA, Goans C, Loh R, Ryall K, Middleton MCA, Dalton A (2021): Suicidal ideation is insensitive to suicide risk after emergency department discharge: Performance characteristics of the Columbia-Suicide Severity Rating Scale Screener. *Acad Emerg Med* 28: 621–629.
19. Simon GE, Rutter CM, Peterson D, Oliver M, Whiteside U, Operskalski B, Ludman EJ (2013): Does response on the PHQ-9 Depression Questionnaire predict subsequent suicide attempt or suicide death? *Psychiatr Serv* 64: 1195–1202.
20. Simon GE, Coleman KJ, Rossom RC, Beck A, Oliver M, Johnson E, *et al.* (2016): Risk of suicide attempt and suicide death following completion of the Patient Health Questionnaire depression module in community practice. *J Clin Psychiatry* 77: 221–227.
21. Kessler RC, Ustün TB (2004): The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res* 13: 93–121.
22. Nurnberger JI, Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

J, *et al.* (1994): Diagnostic Interview for Genetic Studies: Rationale, Unique Features, and Training. *Arch Gen Psychiatry* 51: 849–859.

23. First MB (2015): Structured Clinical Interview for theDSM(SCID). *The Encyclopedia of Clinical Psychology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp 1–6.
24. Horowitz LM, Bridge JA, Teach SJ, Ballard E, Klima J, Rosenstein DL, *et al.* (2012): Ask Suicide-Screening Questions (ASQ): a brief instrument for the pediatric emergency department. *Arch Pediatr Adolesc Med* 166: 1170–1176.
25. Osman A, Bagge CL, Gutierrez PM, Konick LC, Kopper BA, Barrios FX (2001): The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment* 8: 443–454.
26. Kroenke K, Spitzer RL, Williams JBW, Löwe B (2010): The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *Gen Hosp Psychiatry* 32: 345–359.
27. Na PJ, Yaramala SR, Kim JA, Kim H, Goes FS, Zandi PP, *et al.* (2018): The PHQ-9 Item 9 based screening for suicide risk: a validation study of the Patient Health Questionnaire (PHQ)−9 Item 9 with the Columbia Suicide Severity Rating Scale (C-SSRS). *J Affect Disord* 232: 34–40.
28. Thom R, Hogan C, Hazen E (2020): Suicide Risk Screening in the Hospital Setting: A Review of Brief Validated Tools. *Psychosomatics* 61: 1–7.
29. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, *et al.* (2011): The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 168: 1266–1277.
30. Snowdon J, Choi NG (2020): Undercounting of suicides: Where suicide data lie hidden. *Glob Public Health* 15: 1894–1901.
31. Kapusta ND, Tran US, Rockett IRH, De Leo D, Naylor CPE, Niederkrotenthaler T, *et al.* (2011): Declining autopsy rates and suicide misclassification: a cross-national analysis of

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

35 countries. *Arch Gen Psychiatry* 68: 1050–1057.

32. Bakst SS, Braun T, Zucker I, Amitai Z, Shohat T (2016): The accuracy of suicide statistics: are true suicide deaths misclassified? *Soc Psychiatry Psychiatr Epidemiol* 51: 115–123.

33. Kinderman P, Allsopp K, Zero R, Handerer F, Tai S (2021): Minimal use of ICD social determinant or phenomenological codes in mental health care records. *J Ment Health* 1–10.

34. Haerian K, Salmasian H, Friedman C (2012): Methods for identifying suicide or suicidal ideation in EHRs. *AMIA Annu Symp Proc* 2012: 1244–1253.

35. Bejan CA, Ripperger M, Wilimitis D, Ahmed R, Kang J, Robinson K, *et al.* (2022): Improving ascertainment of suicidal ideation and suicide attempt with natural language processing. *Sci Rep* 12: 15146.

36. Zhong Q-Y, Karlson EW, Gelaye B, Finan S, Avillach P, Smoller JW, *et al.* (2018): Screening pregnant women for suicidal behavior in electronic medical records: diagnostic codes vs. clinical notes processed by natural language processing. *BMC Med Inform Decis Mak* 18: 30.

37. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM (2005): Measuring diagnoses: ICD code accuracy. *Health Serv Res* 40: 1620–1639.

38. Hedegaard H, Schoenbaum M, Claassen C, Crosby A, Holland K, Proescholdbell S (2018): Issues in Developing a Surveillance Case Definition for Nonfatal Suicide Attempt and Intentional Self-harm Using International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Coded Data. *Natl Health Stat Report* 1–19.

39. Bigdeli TB, Fanous AH, Li Y, Rajeevan N, Sayward F, Genovese G, *et al.* (2021): Genome-Wide Association Studies of Schizophrenia and Bipolar Disorder in a Diverse Cohort of US Veterans. *Schizophr Bull* 47: 517–529.

40. Bigdeli TB, Barr PB, Rajeevan N, Graham DP, Li Y, Meyers JL, *et al.* (2024): Correlates of suicidal behaviors and genetic risk among United States veterans with schizophrenia or bipolar I disorder. *Mol Psychiatry*. https://doi.org/10.1038/s41380-024-02472-1

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

41. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, Masys DR (2008): Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther* 84: 362–369.

42. Bachmann S (2018): Epidemiology of Suicide and the Psychiatric Perspective. *Int J Environ Res Public Health* 15. https://doi.org/10.3390/ijerph15071425

43. Centers for Disease Control and Prevention (2023, May 8): Facts About Suicide. *Facts About Suicide*. Retrieved February 20, 2024, from https://www.cdc.gov/suicide/facts/index.html

44. Nock MK, Borges G, Bromet EJ, Alonso J, Angermeyer M, Beautrais A, *et al.* (2008): Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br J Psychiatry* 192: 98–105.

45. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, Silove D (2014): The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int J Epidemiol* 43: 476–493.

46. Beautrais AL, Joyce PR, Mulder RT, Fergusson DM, Deavoll BJ, Nightingale SK (1996): Prevalence and comorbidity of mental disorders in persons making serious suicide attempts: a case-control study. *Am J Psychiatry* 153: 1009–1014.

47. van Rheenen W, Peyrot WJ, Schork AJ, Lee SH, Wray NR (2019): Genetic correlations of polygenic disease traits: from theory to practice. *Nat Rev Genet* 20: 567–581.

48. Cheung KL, Ten Klooster PM, Smit C, de Vries H, Pieterse ME (2017): The impact of non-response bias due to sampling in public health studies: A comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC Public Health* 17: 276.

49. Preacher KJ, Rucker DD, MacCallum RC, Nicewander WA (2005): Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychol Methods* 10: 178–192.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

50. Uher R, Rutter M (2012): Basing psychiatric classification on scientific foundation: problems and prospects. *Int Rev Psychiatry* 24: 591–605.

51. Kendler KS, Chatzinakos C, Bacanu S-A (2020): The impact on estimations of genetic correlations by the use of super-normal, unscreened, and family-history screened controls in genome wide case-control studies. *Genet Epidemiol* 44: 283–289.

52. Carry PM, Vanderlinden LA, Dong F, Buckner T, Litkowski E, Vigers T, *et al.* (2021): Inverse probability weighting is an effective method to address selection bias during the analysis of high dimensional data. *Genet Epidemiol* 45: 593–603.

53. Glenn CR, Klonsky ED (2009): Social context during non-suicidal self-injury indicates suicide risk. *Pers Individ Dif* 46: 25–29.

54. Taylor PJ, Jomar K, Dhingra K, Forrester R, Shahmalak U, Dickson JM (2018): A meta-analysis of the prevalence of different functions of non-suicidal self-injury. *J Affect Disord* 227: 759–769.

55. Zahl DL, Hawton K (2004): Repetition of deliberate self-harm and subsequent suicide risk: long-term follow-up study of 11,583 patients. *Br J Psychiatry* 185: 70–75.

56. Corcoran P, Reulbach U, Perry IJ, Arensman E (2010): Suicide and deliberate self harm in older Irish adults. *Int Psychogeriatr* 22: 1327–1336.

57. Lloyd-Richardson EE, Perrine N, Dierker L, Kelley ML (2007): Characteristics and functions of non-suicidal self-injury in a community sample of adolescents. *Psychol Med* 37: 1183–1192.

58. Kimbrel NA, Thomas SP, Hicks TA, Hertzberg MA, Clancy CP, Elbogen EB, *et al.* (2018): Wall/object punching: An important but under-recognized form of nonsuicidal self-injury. *Suicide Life Threat Behav* 48: 501–511.

59. Carroll R, Thomas KH, Bramley K, Williams S, Griffin L, Potokar J, Gunnell D (2016): Self-cutting and risk of subsequent suicide. *J Affect Disord* 192: 8–10.

60. Fernando DT, Clapperton A, Berecki-Gisolf J (2022): Suicide following hospital admission

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

for mental health conditions, physical illness, injury and intentional self-harm in Victoria, Australia. *PLoS One* 17: e0271341.

61. Nada-Raja S, Skegg K, Langley J, Morrison D, Sowerby P (2004): Self-harmful behaviors in a population-based sample of young adults. *Suicide Life Threat Behav* 34: 177–186.
62. Johansen NJ, Christensen MB (2018): A Systematic Review on Insulin Overdose Cases: Clinical Course, Complications and Treatment Options. *Basic Clin Pharmacol Toxicol* 122: 650–659.
63. von Mach MA, Meyer S, Omogbehin B, Kann PH, Weilemann LS (2004): Epidemiological assessment of 160 cases of insulin overdose recorded in a regional poisons unit. *Int J Clin Pharmacol Ther* 42: 277–280.
64. Austin AE, Proescholdbell SK, Creppage KE, Asbun A (2017): Characteristics of self-inflicted drug overdose deaths in North Carolina. *Drug Alcohol Depend* 181: 44–49.
65. Hsieh W-H, Wang C-H, Lu T-H (2018): Drowning mortality by intent: a population-based cross-sectional study of 32 OECD countries, 2012-2014. *BMJ Open* 8: e021501.
66. Lohner J, Konrad N (2006): Deliberate self-harm and suicide attempt in custody: distinguishing features in male inmates' self-injurious behavior. *Int J Law Psychiatry* 29: 370–385.
67. Randall JR, Roos LL, Lix LM, Katz LY, Bolton JM (2017): Emergency department and inpatient coding for self-harm and suicide attempts: Validation using clinician assessment data. *Int J Methods Psychiatr Res* 26. https://doi.org/10.1002/mpr.1559
68. Cai N, Revez JA, Adams MJ, Andlauer TFM, Breen G, Byrne EM, *et al.* (2020): Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat Genet* 52: 437–447.
69. Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, *et al.* (2016): Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* 48: 1031–1036.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

70. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, *et al.* (2018): Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 50: 668–681.

71. Huang X, Ribeiro JD, Franklin JC (2020): The Differences Between Individuals Engaging in Nonsuicidal Self-Injury and Suicide Attempt Are Complex (vs. Complicated or Simple). *Front Psychiatry* 11: 239.

72. van der Ven E, Olino TM, Diehl K, Nuñez SM, Thayer G, Bridgwater MA, *et al.* (2024): Ethnoracial Risk Variation Across the Psychosis Continuum in the US: A Systematic Review and Meta-Analysis. *JAMA Psychiatry* 81: 447–455.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure Legends

Figure 1: Schematic of comparison between SA cases versus potential control groups

A-C. The left panels represent SA cases and the right panels represent the control group. Amongst SA cases, the prevalence of psychiatric disorders is 0.9. **A.** The population control group displays psychiatric disorders at a prevalence of 0.3 and SA at a prevalence of 0.02. **B.** The SA-screened control group displays psychiatric disorders at a prevalence of 0.3. **C.** The non-psychiatric control group assumes a prevalence of 0 for both psychiatric disorders and SA.

Figure 2: Schematic example of sample processing. Diagram representing an example processing of a control sample for a suicide phenotype study. Processing of the control samples varies based on concerns about control quality, available data, and separate control samples. Controls with phenotypic data should be screened for SI/SA/SD/NSSI. If multiple control sets are available and there are quality concerns, consider leave-one-out analyses. If some control sets are considered unbiased, perform tiered analyses, starting with the most reliable. For a single control sample with bias concerns, use inverse probability weighting based on selection probability for comparison groups.

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Tables

Table 1: Definitions of suicide and self-harm phenotypes

Phenotype (Abbreviation)	Phenotype Definition	Aggregated Phenotype (Abbreviation)		
Suicidal ideation (SI) [a]	Thoughts of engaging in suicide-related behavior (1).	Suicidality/ Suicidal thoughts and behaviors (STBs)		
Suicide attempt (SA) [b]	A non-fatal self-directed potentially injurious behavior with any intent to die as a result of the behavior. A suicide attempt may or may not result in injury (1).		Suicidal behavior (SB)^	Deliberate self-harm (DSH)^
Suicide death (SD) [b]	Death caused by self-directed injurious behavior with any intent to die as a result of the behavior (1).			
Non-suicidal self-injury (NSSI) [b]	The intentional self-inflicted destruction of body tissue without suicidal intention and for purposes not socially sanctioned (2)			

[a] Thoughts [b] Behaviors

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2: Suicide attempt (SA) definition ICD code list

ICD Version	Code_Group_Description	Codes
9_CM	Attempt via poison or toxic ingestion	E950.0-E950.9
9_CM	Attempt via poison or toxic Inhalation	E951.0-E951.8;E952.0-952.9
9_CM	Attempt via asphyxiation or drowning	E953.0-E953.9;E954
9_CM	Attempt via firearm	E955.0-E955.9
9_CM	Attempt via jumping, crashing, or exposure	E957.0-E957.9;E958.0;E958.1;E958.3-E958.9
9_CM	Late effects of self-inflicted injuries	E959
9_CM	Undetermined intent, asphyxiation/drowning	E983.0-E983.9;E984
9_CM	Undetermined intent, firearm injury	E985.0-E985.5
9_CM	Undetermined intent, fall, jumping, crash, fire injury	E987.0-E987.9;E988.0-E988.1;E988.5-E988.6
10_CM	Suicide attempt	T14.91;*A,*D,*S
10_CM	Intentional Poisoning via medications, elements, compounds, venom, and other ingested or applied agents	T36-T65;**2,**2A,**2D,**2S
10_CM	Undetermined intent, poisoning by insulin and oral hypoglycemic agents	T38.3;*4A,*4D,*4S
10_CM	Intentional asphyxiation	T71;**2,**2A,**2D,**2S
10_CM	Undetermined intent, injury by hanging	T71.16;4A,4D,4S
10	Intentional poisoning	X60-X69
10	Intentional asphyxiation	X70
10_CM	Intentional drowning	X71;**XA,**XD,**XS
10_CM	Intentional firearm injury	X72-X74;**XA,**XD,**XS
10_CM	Intentional explosive or fire injury	X75-X76;**XA,**XD,**XS
10_CM	Intentional injury by knife, dagger, or sword	X78.1-X78.2;*XA,*XD,*XS
10_CM	Intentional injury by jumping, crashing, electrocution, or exposure	X80-X83;**XA,**XD,**XS
10	Undetermined intent, hanging, drowning, strangulation, suffocation	Y20, Y21
10_CM	Undetermined intent, drowning	Y21;**XA,**XD,**XS
10_CM	Undetermined intent, firearm injury	Y22-Y24;**XA,**XD,**XS
10	Undetermined intent, firearm injury	Y22-Y24
10_CM	Undetermined intent, explosives or fire injury	Y25-Y26;**XA,**XD,**XS
10	Undetermined intent, explosives or fire injury	Y25-26
10_CM	Undetermined intent, falling, jumping, or crashing	Y30-Y32;**XA,**XD,**XS
10	Undetermined intent, falling, jumping, or crashing	Y30-Y32
10	Sequelae of intentional self-harm	Y87
10_CM	Personal history of suicidal behavior (attempt)	Z91.51

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Key: * = wildcard placeholder for all alphanumeric values, inclusive, used in ICD-10 code specifiers; all ICD-10 codes presume inclusion of the base code prior to the semi-colon (without specifiers)

Table 3: Suicide attempt (SA) screening ICD code list

ICD Version	Code_Group_Description	Codes
9_CM	Attempt via poison or toxic ingestion	E950.0-E950.9
9_CM	Attempt via poison or toxic Inhalation	E951.0-E951.8;E952.0-952.9
9_CM	Attempt via asphyxiation or drowning	E953.0-E953.9;E954
9_CM	Attempt via firearm	E955.0-E955.9
9_CM	Attempt via jumping, crashing, or exposure	E957.0-E957.9;E958.0;E958.1;E958.3-E958.9
9_CM	Late effects of self-inflicted injuries	E959
9_CM	Undetermined intent, asphyxiation/drowning	E983.0-E983.9;E984
9_CM	Undetermined intent, firearm injury	E985.0-E985.7
9_CM	Undetermined intent, fall, jumping, crash, exposure	E987.0-E987.9;E988.0-E988.1;E988.3-E988.7
9_CM	**Accidental fall from building**	**E882**
9_CM	**Accidental injury from firearm**	**E922.0-E922.9**
9_CM	**Self-inflicted injury via cutting/piercing instrument**	**E956**
9_CM	**Self-inflicted injury via scald**	**E958.2**
9_CM	**Undetermined intent, poisoning**	**E980.0-E980.9;E981.0-E981.8;E982.1-E982.9**
9_CM	**Undetermined intent, cutting**	**E986**
9_CM	**Undetermined intent, scald**	**E988.2**
9_CM	**Undetermined intent, other or undefined method**	**E988.8-E988.9**
9_CM	**Undetermined intent, late effects of injury**	**E989**
10_CM	**Nonsuicidal self-harm**	**R45.88**
10_CM	Suicide attempt	T14.91;*A,*D,*S
10_CM	Intentional or **undetermined** poisoning via medications, elements, compounds, venom, and other ingested or applied agents	T36-T65;**2,**2A,**2D,**2S,****4A,**4D,**4S**
10_CM	Intentional or **undetermined** asphyxiation	T71;**2,**2A,**2D,**2S,****4A,**4D,**4S**
10_CM	Undetermined and **accidental** intent, injury by hanging	T71.16;4A, 4D, 4S, **1A,1D,1S**

medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

10_CM	**Accidental fall from building**	**W13.4, W13.8, W13.9; 4**A, 4**D, 4**S**
10_CM	**Accidental discharge of firearm**	**W32.0, W33.0, W34.0; **A, **D, **S**
10	Intentional poisoning	X60-X69
10	Intentional asphyxiation	X70
10_CM	Intentional drowning	X71;**XA,**XD,**XS
10_CM	Intentional firearm injury	X72-X74;**XA,**XD,**XS
10_CM	Intentional explosive or fire injury	X75-X76;**XA,**XD,**XS
10_CM	**Intentional injury by hot object/substance**	**X77;**XA,**XD,**XS**
10_CM	**Intentional injury by any sharp object**	**X78;*XA,*XD,*XS**
10_CM	**Intentional injury by blunt object**	**X79;**XA,**XD,**XS**
10_CM	Intentional injury by jumping, crashing, electrocution, or exposure	X80-X83;**XA,**XD,**XS
10	**Intentional harm by unspecified means**	**X84**
10	**Poisoning, various ingested/inhaled agents**	**Y10-Y21**
10_CM	Undetermined intent, drowning	Y21;**XA,**XD,**XS
10	Undetermined intent, firearm injury	Y22-Y24
10_CM	Undetermined intent, firearm injury	Y22-Y24;**XA,**XD,**XS
10	Undetermined intent, explosives or fire injury	Y25-26
10_CM	Undetermined intent, explosives or fire injury	Y25-Y26;**XA,**XD,**XS
10_CM	Undetermined intent, falling, jumping, or crashing	Y30-Y32;**XA,**XD,**XS
10	Sequelae of intentional self-harm	Y87
10	**Undetermined intent, injury by hot object/substance**	**Y27**
10_CM	**Undetermined intent, injury by hot object/substance**	**Y27;**XA,**XD,**XS**
10	**Undetermined intent, injury by sharp object**	**Y28**
10_CM	**Undetermined intent, injury by sharp object**	**Y28;**XA,**XD,**XS**
10	**Undetermined intent, injury by blunt object**	**Y29**
10_CM	**Undetermined intent, injury by blunt object**	**Y29;**XA,**XD,**XS**
10	Undetermined intent, falling, jumping, or crashing	Y30-Y32
10_CM	**Suicide attempt, alleged or ruled out**	**Z03.89**
10_CM	**Personal history of self-harm**	**Z91.5**
10_CM	Personal history of suicidal behavior (attempt)	Z91.51
10_CM	**Personal history of nonsuicidal self-harm**	**Z91.52**

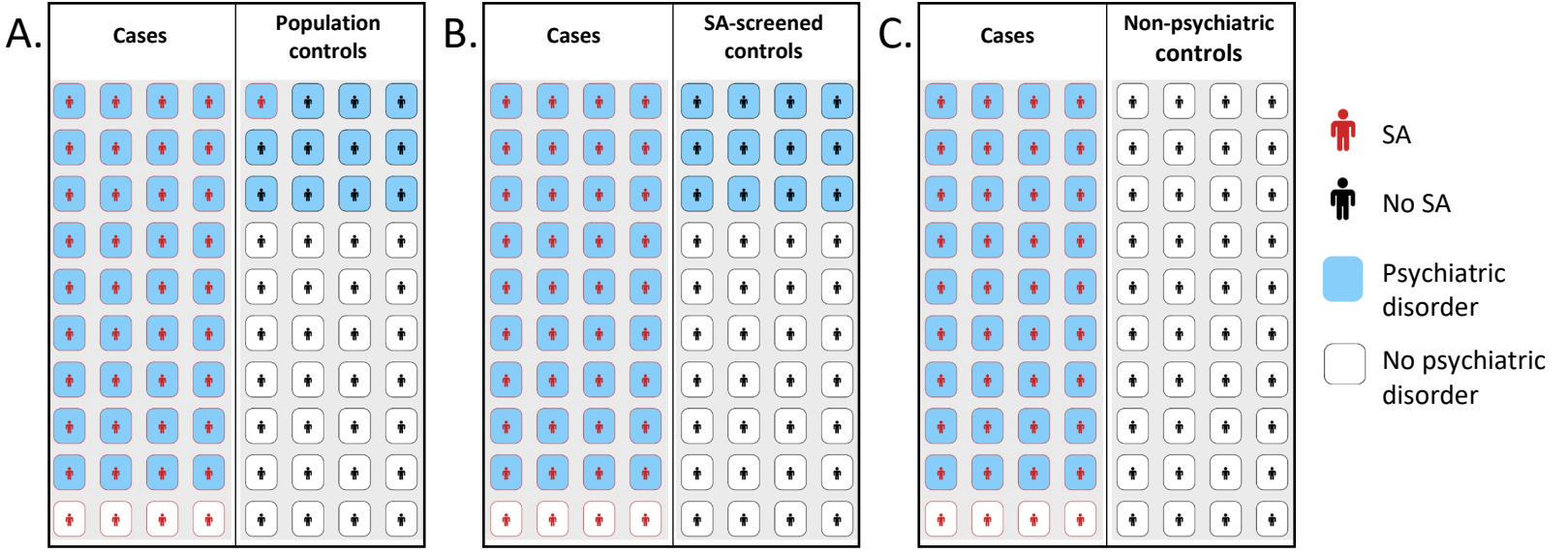
Key: * = wildcard placeholder for alphanumeric values used in all ICD-10 code specifiers, inclusive; all ICD-10 codes presume inclusion of the base code prior to the semi-colon (without specifiers). Bolded codes are those present only in the screening tier and unbolded codes are included in both the definition and screening tiers.

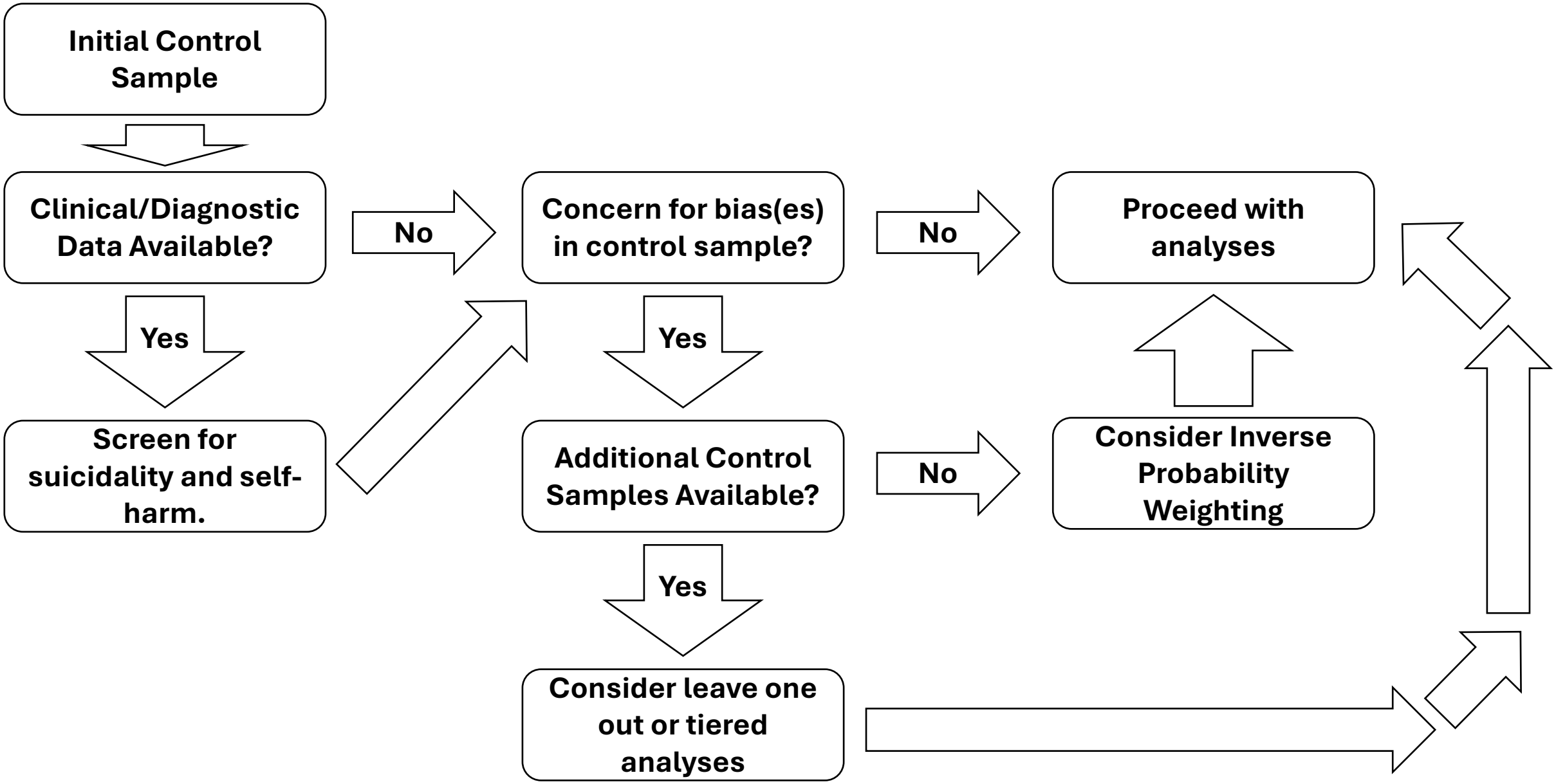
medRxiv preprint doi: https://doi.org/10.1101/2024.07.27.24311110; this version posted July 29, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 4: Suicide Attempt Definition and Screening List Validation Results Versus Instrument Responses in Two Populations

ICD List	Validation Cohort	TP N	FP N	TN N	FN N	SN	SP	PPV
NCHS List	CSP572, BE	710	149	4330	3962	15.20%	96.67%	0.827
SA Definition	CSP572, BE	718	165	4314	3954	15.37%	96.32%	0.813
SA Screen	CSP572, BE	796	218	4261	3876	17.04%	95.13%	0.785
NCHS List	CSP572, AT	1072	310	4169	3600	22.95%	93.08%	0.776
SA Definition	CSP572, AT	1096	353	4126	3576	23.46%	92.12%	0.756
SA Screen	CSP572, AT	1196	423	4056	3476	25.60%	90.56%	0.739
NCHS List	VUMC, All	2339	979	6320	2456	48.78%	86.59%	0.705
SA Definition	VUMC, All	1585	544	6755	3219	32.99%	92.55%	0.744
SA Screen	VUMC, All	2740	1895	5404	2064	57.04%	74.04%	0.591
NCHS List	VUMC, MSA	471	57	322	245	65.78%	84.96%	0.892
SA Definition	VUMC, MSA	415	36	343	301	57.96%	90.50%	0.920
SA Screen	VUMC, MSA	509	123	256	207	71.09%	67.55%	0.805
NCHS List	VUMC, Clin	2035	971	6057	2236	47.65%	86.18%	0.677
SA Definition	VUMC, Clin	1311	543	6485	2960	30.70%	92.27%	0.707
SA Screen	VUMC, Clin	2403	1836	5192	1868	56.26%	73.88%	0.567

Key: TP = true positive; FP = false positive; TN = True negative; FN = False negative; SN = Sensitivity; N = counted total number of individuals; SP = Specificity; PPV = Positive Predictive Value; SA = Suicide attempt; NCHS List = National Center for Health Statistics published ICD list; BE = Before study enrollment; AT = any time (after study enrollment); VUMC = Vanderbilt University Medical Center; MSA = Manually selected individuals with a high enrichment of prior history of suicide attempt; Clin = Psychiatric urgent care clinical sample.


A.
Cases
Population controls
B.
Cases
SA-screened controls
C.
Cases
Non-psychiatric controls
SA
No SA
Psychiatric disorder
No psychiatric disorder


Initial Control Sample
Clinical/Diagnostic Data Available?
No
Yes
Screen for suicidality and self-harm.
Concern for bias(es) in control sample?
No
Yes
Additional Control Samples Available?
No
Yes
Consider leave one out or tiered analyses
Proceed with analyses
Consider Inverse Probability Weighting