



## Original Research

## Global cross-cultural validation of a brief measure for identifying potential suicide risk in 42 countries



A. Gewirtz-Meydan<sup>a,\*</sup>, M. Koós<sup>c,bx</sup>, L. Nagy<sup>b,c</sup>, S.W. Kraus<sup>d</sup>, Z. Demetrovics<sup>c,e</sup>, M.N. Potenza<sup>f,g</sup>, R. Ballester-Arnal<sup>h</sup>, D. Batthyány<sup>i</sup>, S. Bergeron<sup>j</sup>, J. Billieux<sup>k,l</sup>, J. Burkauskas<sup>m</sup>, G. Cárdenas-López<sup>n</sup>, J. Carvalho<sup>o</sup>, J. Castro-Calvo<sup>p</sup>, L. Chen<sup>q</sup>, G. Ciocca<sup>r</sup>, O. Corazza<sup>s,t</sup>, R. Csako<sup>u</sup>, D.P. Fernandez<sup>v</sup>, H. Fujiwara<sup>w,x</sup>, E.F. Fernandez<sup>y</sup>, J. Fuss<sup>z</sup>, R. Gabrhelík<sup>aa,ab</sup>, B. Gjonjeska<sup>ac</sup>, M. Gola<sup>ad,ae</sup>, J.B. Grubbs<sup>af,ag</sup>, H.T. Hashim<sup>ah</sup>, M.S. Islam<sup>ai,aj</sup>, M. Ismail<sup>ah</sup>, M.C. Jiménez-Martínez<sup>ak,al</sup>, T. Jurin<sup>am</sup>, O. Kalina<sup>an</sup>, V. Klein<sup>ao</sup>, A. Költő<sup>ap</sup>, S.-K. Lee<sup>aq,ar</sup>, K. Lewczuk<sup>as</sup>, C.-Y. Lin<sup>at,au</sup>, C. Lochner<sup>av</sup>, S. López-Alvarado<sup>aw</sup>, K. Lukavská<sup>aa,ax</sup>, P. Mayta-Tristán<sup>ay</sup>, D.J. Miller<sup>az</sup>, O. Orosová<sup>ba</sup>, G. Orosz<sup>bb</sup>, Sungkyunkwan University's research team NA<sup>bc,\*</sup>, F.P. Ponce<sup>bd</sup>, G.R. Quintana<sup>be</sup>, G.C. Quintero Garzola<sup>bf,bg</sup>, J. Ramos-Díaz<sup>bh</sup>, K. Rigaud<sup>bb</sup>, A. Rousseau<sup>bi</sup>, M. De Tubino Scanavino<sup>bj,bk</sup>, M.K. Schulmeyer<sup>bl</sup>, P. Sharan<sup>bm</sup>, M. Shibata<sup>w</sup>, S. Shoib<sup>bn,bo</sup>, V. Sigre-Leirós<sup>k,bp</sup>, L. Sniewski<sup>bq</sup>, O. Spasovski<sup>br,bs</sup>, V. Steibliene<sup>m</sup>, D.J. Stein<sup>bt</sup>, C. Strong<sup>bu</sup>, B.C. Ünsal<sup>b,c</sup>, M.-P. Vaillancourt-Morel<sup>bv</sup>, M.C. Van Hout<sup>bw</sup>, B. Bóthe<sup>j,bv</sup>

<sup>a</sup> School of Social Work, Faculty of Social Welfare and Health Sciences, University of Haifa, Israel

<sup>b</sup> Doctoral School of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

<sup>c</sup> Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

<sup>d</sup> Department of Psychology, University of Nevada, Las Vegas, Las Vegas, NV, USA

<sup>e</sup> Centre of Excellence in Responsible Gaming, University of Gibraltar, Gibraltar, Gibraltar

<sup>f</sup> Yale University School of Medicine, New Haven, CT, USA

<sup>g</sup> Connecticut Council on Problem Gambling, Wethersfield, CT, USA

<sup>h</sup> Departamento de Psicología Básica, Clínica y Psicobiología, University Jaume I of Castellón, Spain

<sup>i</sup> Institute for Behavioural Addictions, Sigmund Freud University Vienna, Austria

<sup>j</sup> Département de Psychologie, Université de Montréal, Montréal, Canada

<sup>k</sup> Institute of Psychology, University of Lausanne, Lausanne, Switzerland

<sup>l</sup> Center for Excessive Gambling, Addiction Medicine, Lausanne University Hospitals (CHUV), Lausanne, Switzerland

<sup>m</sup> Laboratory of Behavioral Medicine, Neuroscience Institute, Lithuanian University of Health Sciences, Lithuania

<sup>n</sup> Virtual Teaching and Cyberpsychology Laboratory, School of Psychology, National Autonomous University of Mexico, Mexico

<sup>o</sup> William James Center for Research, Departamento de Educação e Psicologia, Universidade de Aveiro, Aveiro, Portugal

<sup>p</sup> Department of Personality, Assessment, and Psychological Treatments, University of Valencia, Spain

<sup>q</sup> Department of Psychology, College of Humanity and Social Science, Fuzhou University, China

<sup>r</sup> Section of Sexual Psychopathology, Department of Dynamic and Clinical Psychology, and Health Studies, Sapienza University of Rome, Rome, Italy

<sup>s</sup> Department of Clinical, Pharmaceutical and Biological Sciences, University of Hertfordshire, UK

<sup>t</sup> Department of Psychology and Cognitive Science, University of Trento, Italy

<sup>u</sup> Department of Psychology and Neuroscience, Auckland University of Technology, Auckland, New Zealand

<sup>v</sup> Nottingham Trent University, UK

<sup>w</sup> Department of Neuropsychiatry, Graduate School of Medicine, Kyoto University, Kyoto, Japan

<sup>x</sup> Decentralized Big Data Team, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

<sup>y</sup> HELP University, Malaysia

<sup>z</sup> Institute of Forensic Psychiatry and Sex Research, Center for Translational Neuro- and Behavioral Sciences, University of Duisburg-Essen, Essen, Germany

<sup>aa</sup> Charles University, First Faculty of Medicine, Department of Addictology, Prague, Czech Republic

<sup>ab</sup> General University Hospital in Prague, Department of Addictology, Czech Republic

<sup>ac</sup> Macedonian Academy of Sciences and Arts, Macedonia

<sup>ad</sup> Institute of Psychology, Polish Academy of Sciences, Poland

<sup>ae</sup> Institute for Neural Computations, University of California San Diego, USA

\* Corresponding author. School of Social Work, Faculty of Social Welfare and Health Sciences, University of Haifa, 199 Abba Khoushy, Mount Carmel, Haifa, 3498838, Israel. Tel.: +9724 824-0812.

E-mail address: [Agewirtz@univ.haifa.ac.il](mailto:Agewirtz@univ.haifa.ac.il) (A. Gewirtz-Meydan).

\* The Sungkyunkwan University's research team includes Dr. H. Chang and Mr. K. Park.

- <sup>af</sup> University of New Mexico, Albuquerque, USA  
<sup>ag</sup> Center for Alcohol, Substance Use, And Addiction (CASAA), University of New Mexico, Albuquerque, USA  
<sup>ah</sup> University of Baghdad, College of Medicine, Iraq  
<sup>ai</sup> Department of Public Health and Informatics, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh  
<sup>aj</sup> Centre for Advanced Research Excellence in Public Health, Savar, Dhaka 1342, Bangladesh  
<sup>ak</sup> Universidad Pedagógica y Tecnológica de Colombia, Colombia  
<sup>al</sup> Grupo de Investigación Biomédica y de Patología, Colombia  
<sup>am</sup> Department of Psychology, Humanities and Social Sciences, University of Zagreb, Croatia  
<sup>an</sup> Department of Educational Psychology and Psychology of Health, Pavol Jozef Safarik University in Kosice, Slovakia  
<sup>ao</sup> School of Psychology, University of Southampton, UK  
<sup>ap</sup> Health Promotion Research Centre, University of Galway, Ireland, UK  
<sup>aq</sup> Department of Psychiatry, Hallym University Chuncheon Sacred Heart Hospital, South Korea  
<sup>ar</sup> Chuncheon Addiction Management Center, South Korea  
<sup>as</sup> Institute of Psychology, Cardinal Stefan Wyszyński University, Warsaw, Poland  
<sup>at</sup> Institute of Allied Health Sciences, College of Medicine, National Cheng Kung University, Tainan, Taiwan  
<sup>au</sup> Biostatistics Consulting Center, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan  
<sup>av</sup> SAMRC Unit on Risk & Resilience in Mental Disorders, Stellenbosch University, South Africa  
<sup>aw</sup> Faculty of Psychology, University of Cuenca, Cuenca, Ecuador  
<sup>ax</sup> Charles University, Faculty of Education, Department of Psychology, Prague, Czech Republic  
<sup>ay</sup> Facultad de Medicina, Universidad Científica del Sur, Lima, Peru  
<sup>az</sup> College of Healthcare Sciences, James Cook University, Australia  
<sup>ba</sup> Pavol Jozef Safarik University in Kosice, Department of Educational Psychology and Psychology of Health, Slovakia  
<sup>bb</sup> Artois University, France  
<sup>bc</sup> Department of Psychology, Sungkyunkwan University, South Korea  
<sup>bd</sup> Facultad de Psicología, Universidad de Talca, Chile  
<sup>be</sup> Departamento de Psicología y Filosofía, Facultad de Ciencias Sociales, Universidad de Tarapacá, Arica, Arica y Parinacota, Chile  
<sup>bf</sup> Florida State University, Panama  
<sup>bg</sup> Sistema Nacional de Investigación (SNI), SENACYT, Panama  
<sup>bh</sup> Facultad de Ciencias de la Salud, Universidad Privada del Norte, Lima, Perú  
<sup>bi</sup> Leuven School for Mass Communication, KU Leuven, Leuven, Belgium  
<sup>bj</sup> Department of Psychiatry, Faculdade de Medicina, Universidade de São Paulo, Brazil  
<sup>bk</sup> Experimental Pathophysiology Post Graduation Program, Faculdade de Medicina, Universidade de São Paulo, Brazil  
<sup>bl</sup> Universidad Privada de Santa Cruz de la Sierra, Bolivia  
<sup>bm</sup> Department of Psychiatry, All India Institute of Medical Sciences, New Delhi, 110029, India  
<sup>bn</sup> Department of Psychology, Sharda University, India  
<sup>bo</sup> Department of Health Kashmir, India  
<sup>bp</sup> Institute of Legal Psychiatry, Lausanne University Hospitals (CHUV), Lausanne, Switzerland  
<sup>bq</sup> Auckland University of Technology, New Zealand  
<sup>br</sup> Faculty of Philosophy, Ss. Cyril and Methodius University in Skopje, Macedonia  
<sup>bs</sup> Faculty of Philosophy, University of Ss. Cyril and Methodius in Trnava, Slovakia  
<sup>bt</sup> SAMRC Unit on Risk & Resilience in Mental Disorders, Dept of Psychiatry & Neuroscience Institute, University of Cape Town, South Africa  
<sup>bu</sup> Department of Public Health, College of Medicine, National Cheng Kung University, Tainan, Taiwan  
<sup>bv</sup> Département de Psychologie, Université du Québec à Trois-Rivières, Trois-Rivières, Canada  
<sup>bw</sup> Public Health Institute, Faculty of Health, Liverpool John Moores University, UK  
<sup>bx</sup> Institute of Forensic Psychiatry and Sex Research, University of Duisburg-Essen, Essen, Germany

## ARTICLE INFO

### Article history:

Received 22 August 2023

Received in revised form

11 December 2023

Accepted 29 December 2023

Available online 20 February 2024

### Keywords:

Suicide

Assessment

Validation

Cross-cultural

## ABSTRACT

**Objectives:** This study aimed to examine the psychometric properties of the P4 suicide screener in a multinational sample. The primary goal was to evaluate the reliability and validity of the scale and investigate its convergent validity by analyzing its correlation with depression, anxiety, and substance use.

**Study design:** The study design is a cross-sectional self-report study conducted across 42 countries.

**Methods:** A cross-sectional, self-report study was conducted in 42 countries, with a total of 82,243 participants included in the final data set.

**Results:** The study provides an overview of suicide ideation rates across 42 countries and confirms the structural validity of the P4 screener. The findings indicated that sexual and gender minority individuals exhibited higher rates of suicidal ideation. The P4 screener showed adequate reliability, convergence, and discriminant validity, and a cutoff score of 1 is recommended to identify individuals at risk of suicidal behavior.

**Conclusions:** The study supports the reliability and validity of the P4 suicide screener across 42 diverse countries, highlighting the importance of using a cross-cultural suicide risk assessment to standardize the identification of high-risk individuals and tailoring culturally sensitive suicide prevention strategies.

© 2024 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

## Introduction

Suicide refers to the act of intentionally causing one's own death, with the risk often indicated by suicidal ideation or intent, especially when accompanied by a well-elaborated suicidal plan

(American Psychological Association), and represents a significant societal and healthcare problem. Broader terms encompassing various outcomes related to suicide include 'suicide-related outcomes' and 'suicidal thoughts and behaviors.' These terms cover a spectrum of behaviors, including suicidal ideation (thoughts of self-

harm or death), suicide attempts (engaging in actions with the intent to die), and completed suicide (resulting in the individual's death). Widely used in research and clinical settings, these terms help capture the diverse range of behaviors and thoughts associated with deliberate self-harm and suicide, playing a crucial role in understanding and addressing these complex issues.

According to the World Health Organization,<sup>1</sup> more than 700,000 people die due to suicide every year, placing suicide as the fourth leading cause of death among 15- to 29-year-olds. Multiple predisposing and precipitating factors have been identified as contributing to suicides among adults in the general population,<sup>2,3</sup> with mental health disorders—mood, substance use, and psychotic disorders—being the most important predictors.<sup>4</sup> Although suicide remains a significant public health concern, previous research has focused predominantly on Western, Educated, Industrialized, Rich and Democratic (WEIRD) populations, leaving a considerable knowledge gap regarding how suicidal ideation manifests in diverse populations.<sup>1</sup> This study addresses this gap through a large-scale, international, multilaboratory, multilanguage approach using cross-sectional survey methods to examine suicidal ideation and behavior using the P4 suicide screener across 42 countries.

The P4 screener validly and reliably evaluates past suicide attempts, the presence of a suicidal plan, the likelihood of completion, and the presence of protective factors (e.g. resilience, access to mental health services).<sup>5,6</sup> Developed over multiple controlled trials,<sup>7</sup> the P4 screener possesses adequate clinical validity and was selected for this study due to its culturally sensitive language, which prioritizes ethical considerations in regions with cultural sensitivity toward the topic of suicide.<sup>8</sup> A recent review highlights the varied use of the P4 screener in diverse settings, such as psychiatry inpatients and veterans, as well as its application for safety monitoring in a clinical trial. The P4 has proven to be an efficient tool for categorizing individuals expressing thoughts of deliberate self-harm into ordinal risk categories.<sup>6</sup> The language used in the P4 screener is less explicit regarding the topic of suicide, which was deemed appropriate for the various cultural contexts of the population participating in this study.

### The present study

Validated cross-cultural suicide risk assessment scales are important for accurate and unbiased assessment of suicide risk. These scales standardize and ensure reliability in determining an individual's suicide risk, effectively identify high-risk individuals, and monitor changes in risk levels, leading to better outcomes for those at risk of suicide. The primary aim of this study was to examine the psychometric properties of the P4 suicidality screener in a multinational sample. Our primary goal was to evaluate the reliability and validity of the scale by using appropriate statistical techniques. Second, we aimed to investigate the convergent validity of the P4 suicidality screener by analyzing its correlation with established measures of depression and anxiety levels, and substance use. Moreover, we aimed to investigate variations in rates of suicidal ideation across different countries, genders, and sexual orientations.

## Methods

### Procedures

The International Sex Survey (ISS) is a cross-sectional, self-report study conducted in 42 countries. The study design, including the preregistered study protocol, can be found at [https://osf.io/uyfra/?view\\_only=6e4f96b748be42d99363d58e32d511b8](https://osf.io/uyfra/?view_only=6e4f96b748be42d99363d58e32d511b8). The ISS was initiated in February 2021, following the acquisition of ethical

permissions by collaborators. Although the study aimed to use extensive and varied samples, achieving a comprehensive representation of all populations within each country proved challenging. Nonetheless, collaborators from each country collected a community sample of adults, striving for a balanced gender ratio of approximately 50%–50% for both men and women and representation of diverse individuals with respect to sexuality and gender within the surveyed population. The English survey battery was translated into 25 other languages following a pre-established procedure for cross-cultural studies.<sup>9</sup> Data collection took place between October 2021 and May 2022. Participants aged  $\geq 18$  years (or the legal age of being adult in a given country) were recruited through advertisements and completed the anonymous survey on Qualtrics (a secure online platform), which took approximately 25–45 min. Rigorous attention checks were implemented to ensure the data's reliability, and participants failing these checks were excluded from the study (see the detailed data cleaning procedure at <https://doi.org/10.17605/OSF.IO/DK78R>). Incentives included a donation of 50 cents (USD) to non-profit, sexuality-related international organizations for each completed survey, with a maximum donation of \$1000. Participants could choose their preferred organization from a list after completing the survey. The list of collaborating countries, detailed information on the translation and data collection procedures, and eligibility criteria can be found in the study protocol.<sup>10</sup>

### Participants

After data cleaning (see detailed data cleaning procedure: <https://doi.org/10.17605/OSF.IO/DK78R>), a total of 82,243 participants ( $M_{age} = 32.39$  years,  $SD = 12.52$ ) were included in the final data set. Concerning participants' gender, 32,549 (39.6%) were men, 46,874 (57.0%) were women, and 2783 (3.4%) were gender-diverse individuals. Most participants ( $n = 56,125$ ; 68.2%) identified as heterosexual, while 31.5% ( $n = 25,777$ ) of the participants belonged to sexual minorities. Most participants completed tertiary education (e.g. college or university;  $n = 60,896$ ; 74.0%), worked full time ( $n = 42,981$ ; 52.3%), and lived in a city or metropolis (i.e. in a city with a population greater than 100,000;  $n = 56,361$ ; 68.5%). More than half of the participants were in a romantic relationship ( $n = 51,778$ ; 63.0%). Details on participants' sociodemographic characteristics are presented in Table 1, and the same information disaggregated by country can be found at [https://osf.io/n3k2c/?view\\_only=838146f6027c4e6bb68371d9d14220b5](https://osf.io/n3k2c/?view_only=838146f6027c4e6bb68371d9d14220b5).

### Measures

#### Sociodemographic and sexuality-related questions

Several sociodemographic questions (e.g. gender) and sexuality-related questions were included in the ISS.<sup>11</sup> The translation of the measures in this study can be found at the following link in all available languages: [https://osf.io/jc296/?view\\_only=9af0068dde81488db54638a01c8ae118](https://osf.io/jc296/?view_only=9af0068dde81488db54638a01c8ae118).

Suicide risk was assessed using the P4 screener, which is a brief measure for assessing suicide risk by asking about four key factors: past suicide attempts, a plan for suicide, the probability of completing suicide, and preventive factors that may reduce the likelihood of deliberate self-harm.<sup>7</sup> The P4 screener begins by asking a screening question: 'Have you had thoughts of actually hurting yourself?' (yes/no). If participants answer yes, the screener proceeds with four additional questions: (1) Have you ever attempted to harm yourself in the past? (yes = 1/no = 0); (2) Have you thought about how you might actually hurt yourself? (yes = 1/no = 0); (3) There's a big difference between having a thought and acting on a thought. How likely do you think it is that you will act on

these thoughts about hurting yourself or ending your life sometime over the next month? (not at all likely = 0/somewhat likely = 1/very likely = 2); (4) Is there anything that would prevent or keep you from harming yourself? (yes = 0/no = 1). If respondents chose yes, they were asked to specify what would prevent them from deliberately harming themselves. A sum score was calculated for each participant such that it ranged between 0 and 5 (i.e. item 3 was not collapsed).

*Mental health* was assessed using the two subscales (anxiety and depression) of the Brief Symptom Inventory (BSI),<sup>42</sup> a 12-item scale used to assess anxiety and depression. Cronbach's alpha for both subscales was excellent ( $\alpha = 0.90$ ).<sup>42</sup>

*Substance use* was assessed using the Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST).<sup>13,14</sup> Ten ASSIST items were used to assess the frequency and consequences of substance use in the past 3 months. Cronbach's alpha for the ASSIST scale was acceptable ( $\alpha = 0.63$ ).<sup>43</sup>

*Statistical analyses*

*Descriptive analyses*

As per the preregistered analysis plan (<https://doi.org/10.17605/OSF.IO/DK78R>), we computed descriptive statistics for all items of the P4 screener, which included means with standard deviation, minimum and maximum values, skewness and kurtosis values, and the ratio of 'yes' answers. However, missing values were observed in country, language, gender, and sexual orientation variables (ranging from 0 to 0.24%) and were not missing completely at random. This was supported by the results of Little's Missing Completely at Random Test ( $\chi^2 = 1137.0$ ,  $df = 28$ ,  $P < .001$ ; 15 missing patterns).<sup>15</sup> We excluded participants who had missing data on all items of the P4 screener, allowing partial missingness. To handle the remaining missing data, we used the full-information maximum likelihood method based on previous recommendations.<sup>16</sup>

*Test of dimensionality*

Confirmatory factor analyses (CFAs) were conducted to examine the structural validity and dimensionality of the suicidal ideation scale separately for participants' language, country, gender, and sexual orientation. Based on Monte Carlo simulations (see details: <https://doi.org/10.17605/OSF.IO/DK78R>), a minimum of 535 participants were required to be included in analyses. In the first set of CFAs, participants' language was the grouping variable with 20 languages (see all languages in Table 1), as 20 of the 25 languages had enough participants (i.e.  $n > 535$ ) for these tests. In the second set of analyses, participants' country of residence was the grouping variable with 31 countries (see Table 1 for all countries), as 31 of the 42 countries had sufficient participants for these analyses. In the third set of analyses, participants' gender identity was the grouping variable with three categories (i.e. men, women, gender-diverse individuals), as the number of participants in different gender minority groups did not allow us to use them as separate groups. In the fourth set of analyses, participants' sexual orientation was the grouping variable with eight sexual orientations (see all sexual orientations in Table 1). Information on the creation of gender-based and sexual orientation-based groups can be found in the preregistration document (<https://doi.org/10.17605/OSF.IO/DK78R>).

The CFAs were evaluated using common goodness-of-fit indices:<sup>17–19</sup> Comparative Fit Index ( $\geq 0.90$  adequate;  $\geq 0.95$  good), Tucker–Lewis Index ( $\geq 0.90$  adequate;  $\geq 0.95$  good), Root Mean Square Error of Approximation with its 90% confidence interval (CI;  $\geq 0.10$  acceptable,  $\leq 0.08$  adequate, and  $\leq 0.05$  good), and Standardized Root Mean Square Residual ( $\leq 0.08$  good).<sup>20</sup> The mean and variance-adjusted unweighted least squares was used, which is superior to the maximum likelihood estimation for binary items,

**Table 1**  
Participants' sociodemographic characteristics (N = 73,531).

Characteristic	Values <sup>a</sup>
Country of residence	
Algeria	18 (<0.1%)
Australia	577 (0.8%)
Austria	689 (0.9%)
Bangladesh	328 (0.4%)
Belgium	584 (0.8%)
Bolivia	313 (0.4%)
Brazil	3140 (4.3%)
Canada	2285 (3.1%)
China	2395 (3.3%)
Colombia	1771 (2.4%)
Croatia	2133 (2.9%)
Czech Republic	1618 (2.2%)
Ecuador	233 (0.3%)
Egypt	50 (<0.1%)
France	1527 (2.1%)
Germany	3026 (4.1%)
Gibraltar	47 (<0.1%)
Hungary	9974 (14%)
India	151 (0.2%)
Iraq	84 (0.1%)
Ireland	1435 (2.0%)
Israel	1167 (1.6%)
Italy	2047 (2.8%)
Japan	495 (0.7%)
Lithuania	1840 (2.5%)
Malaysia	1085 (1.5%)
Mexico	1821 (2.5%)
New Zealand	2488 (3.4%)
North Macedonia	1147 (1.6%)
Pakistan	0 (0%)
Panama	279 (0.4%)
Peru	2291 (3.1%)
Poland	8653 (12%)
Portugal	1981 (2.7%)
Slovakia	994 (1.4%)
South Africa	1620 (2.2%)
South Korea	1328 (1.8%)
Spain	2101 (2.9%)
Switzerland	1070 (1.5%)
Taiwan	2603 (3.5%)
Turkey	689 (0.9%)
United Kingdom	1249 (1.7%)
United States of America	2147 (2.9%)
Other	975 (1.3%)
Chile	1083 (1.5%)
Language	
Arabic	121 (0.2%)
Bangla	300 (0.4%)
Croatian	2248 (3.1%)
Czech	1568 (2.1%)
Dutch	469 (0.6%)
English	12,310 (17%)
French	3575 (4.9%)
German	3262 (4.4%)
Hebrew	1147 (1.6%)
Hindi	14 (<0.1%)
Hungarian	9724 (13%)
Italian	2080 (2.8%)
Japanese	409 (0.6%)
Korean	1304 (1.8%)
Lithuanian	1913 (2.6%)
Macedonian	1193 (1.6%)
Mandarin – simplified	2437 (3.3%)
Mandarin – traditional	2618 (3.6%)
Polish	9044 (12%)
Portuguese – Brazil	3207 (4.4%)
Portuguese – Portugal	1984 (2.7%)
Slovak	1901 (2.6%)
Spanish – Latin American	7829 (11%)
Spanish – Spain	2091 (2.8%)
Turkish	718 (1.0%)
Sex assigned at birth	
Male	29,733 (40%)
Female	43,788 (60%)

**Table 1** (continued)

Characteristic	Values <sup>a</sup>
Missing values	10
Gender (original answer options in the survey)	
Masculine/man	29,106 (40%)
Feminine/woman	41,881 (57%)
Indigenous or other cultural gender minority identity (e.g. two-spirit)	151 (0.2%)
Non-binary, gender fluid, or something else (e.g. genderqueer)	2089 (2.8%)
Other (If you wish, tell us how you personally describe your gender):	272 (0.4%)
Missing values	32
Gender (categories used in the analyses)	
Man	29,106 (40%)
Woman	41,881 (57%)
Gender-diverse individual	2512 (3.4%)
Missing values	32
Trans status	
No, I am not a trans person	70,833 (96%)
Yes, I am a trans man	329 (0.4%)
Yes, I am a trans woman	271 (0.4%)
Yes, I am a non-binary trans person	798 (1.1%)
I am questioning my gender identity	1032 (1.4%)
I don't know what it means	248 (0.3%)
Missing values	20
Sexual orientation (original answer options in the survey)	
Heterosexual/straight	50,205 (68%)
Gay or lesbian or homosexual	4156 (5.7%)
Heteroflexible	5538 (7.5%)
Homoflexible	483 (0.7%)
Bisexual	6804 (9.3%)
Queer	851 (1.2%)
Pansexual	1747 (2.4%)
Asexual	978 (1.3%)
I do not know yet or I am currently questioning my sexual orientation	1731 (2.4%)
None of the above, specify:	722 (1.0%)
I don't want to answer	288 (0.4%)
Missing values	28
Sexual orientation (categories used in the analyses)	
Heterosexual	50,205 (69%)
Homosexual	4156 (5.7%)
Bisexual	6804 (9.3%)
Queer and pansexual	2598 (3.5%)
Homo- and hetero-flexible identities	6021 (8.2%)
Asexual	978 (1.3%)
Questioning	1731 (2.4%)
Other	722 (1.0%)
Missing values	316
Highest level of education	
Primary (e.g. elementary school)	879 (1.2%)
Secondary (e.g. high school)	17,943 (24%)
Tertiary (e.g. college or university)	54,693 (74%)
Missing values	16
Currently being in education	
No	44,779 (61%)
Yes, in primary education (e.g. elementary school)	55 (<0.1%)
Yes, in secondary education (e.g. high school)	1315 (1.8%)
Yes, in tertiary education (e.g. college or university)	27,342 (37%)
Missing values	40
Work status	
No	18,764 (26%)
Yes, full time	38,379 (52%)
Yes, part time	10,121 (14%)
Yes, I do odd jobs	6247 (8.5%)
Missing values	20
Socio-economic status	
My life circumstances are among the worst	201 (0.3%)
My life circumstances are much worse than average	693 (0.9%)
My life circumstances are worse than average	3833 (5.2%)
My life circumstances are average	23,908 (33%)
My life circumstances are better than average	28,182 (38%)
My life circumstances are much better than average	13,172 (18%)
My life circumstances are among the best	3534 (4.8%)
Missing values	8
Residence	

(continued on next page)

**Table 1** (continued)

Characteristic	Values <sup>a</sup>
Metropolis (population is over 1 million people)	23,767 (32%)
City (population is between 100,000–999,999 people)	26,721 (36%)
Town (population is between 1000–99,999 people)	18,808 (26%)
Village (population is below 1000 people)	4222 (5.7%)
Missing values	13
Relationship status	
Single	24,709 (34%)
In a relationship	24,271 (33%)
Married or common-law partners	21,871 (30%)
Widow or widower	394 (0.5%)
Divorced	2264 (3.1%)
Missing values	22
Having children	
No	51,655 (70%)
yes, 1	7528 (10%)
yes, 2	9328 (13%)
yes, 3	3476 (4.7%)
yes, 4	921 (1.3%)
yes, 5	262 (0.4%)
yes, 6-9	116 (0.2%)
yes, 10 or more	20 (<0.1%)
Missing values	225
Present religion	
Christian	21,348 (29%)
Buddhist	1445 (2.0%)
Hindu	198 (0.3%)
Muslim	890 (1.2%)
Spiritual but not committed to one religion	10,086 (14%)
I am not religious	35,369 (48%)
Other, please specify:	1958 (2.7%)
Jewish	1133 (1.5%)
Taoist	605 (0.8%)
Confucianist	18 (<0.1%)
Sikh	29 (<0.1%)
Spiritist	381 (0.5%)
Jain	14 (<0.1%)
Missing values	57
Age	33 (13), 18
Missing values	–99
Missing values	13

<sup>a</sup> n (%); mean (SD), range.

particularly when the response categories follow asymmetric thresholds.<sup>21</sup> CFAs were estimated in Mplus 8.8. When the goodness-of-fit was unsatisfactory, additional non-pre-registered analyses were undertaken, and an in-depth exploration was conducted using bootstrap Exploratory Graph Analysis,<sup>22</sup> including item and dimensional stability. These analyses were performed in R using the EGAnet package.<sup>23</sup>

**Group differences**

Differences in suicidal ideation by language, country, gender, and sexual orientation were examined using Kruskal–Wallis tests and eta-squared effect sizes (>0.01 weak, >0.06 medium, >0.14 large) followed by Tukey's Honest Significant Differences post hoc tests.<sup>a</sup>

**Tests of reliability and validity**

Cronbach's alphas and McDonald's omegas were calculated to assess the reliability of the suicidal ideation scale (>0.70 good).<sup>24</sup> In cases of low reliability, we also calculated coefficients H (>0.70 good) to appraise whether the low reliability stems from a low factor loading of one of the items.<sup>25</sup> We adopted the Multi-trait Multimethod analysis for assessing the convergence and discriminant validity of

<sup>a</sup> It was not possible to conduct the preregistered measurement invariance tests due to the characteristics of the P4 screener (e.g. binary answer options).

the P4 screener. The analysis was conducted within the *psy* package with the *mtmm* (Multi-trait Multimethod) function.<sup>26</sup> The Multi-trait Multimethod analysis examines within-cluster and between-cluster correlations such that the unit of analysis is a single item. Good psychometric indices are high within-cluster correlations (e.g. high correlations between all items of the P4 screener), moderate between-cluster correlations between items from similar theoretical constructs (i.e. convergent validity), and weak between-cluster correlations between items from unrelated theoretical constructs (i.e. discriminant validity). In the current study, the convergence validity within the Multi-trait Multimethod analysis was estimated between items of the P4 screener and subscales of the BSI scale assessing depressive and anxiety symptoms; the discriminant analysis was appraised by assessing the between-cluster correlations between items of the suicidal ideation scale and items of substance use (ASSIST scale). Correlations around  $|0.10|$  were considered weak,  $|0.30|$  moderate, and  $|0.50|$  strong.<sup>27</sup>

#### Establishing a cutoff score for the ‘at-risk’ group

To establish a cutoff score, we examined the effectiveness of the suicidal ideation score in differentiating between participants with suspected depression and anxiety (above a T score of 63 in BSI depression and anxiety;  $n = 4908$ ) and those without suspected depression and anxiety ( $n = 61,810$ ). To do so, we calculated the optimal clinical cut-point by bootstrapping the optimal cut-point while maximizing the sensitivity and specificity (i.e. highest Youden’s index: sensitivity + specificity – 1). We also reported the suggested indices of the ‘number needed to diagnose’ (NND)<sup>28</sup>—the number of participants who need to be examined to correctly detect one person with the disorder of interest in a study population of persons with and without the known disorder—‘number needed to misdiagnose’ (NNM)<sup>29</sup>—number of individuals who need to be tested for one to be misdiagnosed by the test—and the ‘likelihood to be diagnosed or misdiagnosed’ (LDM),<sup>30</sup> with LDM higher values ( $>1$ ) suggesting that a test is more likely to diagnose than misdiagnose.

## Results

### Psychometric properties of the suicidal ideation scale

#### Descriptive statistics

Descriptive data for the P4 screener items are presented in [Supplementary Table 1](#). Among all participants, 20% reported attempting to deliberately harm themselves, and 23% reported having thought about specific ways to deliberately harm themselves. Six percent of participants reported that it was somewhat to very likely that they would act on their thoughts, and 9.4% reported that nothing was preventing them from deliberately harming themselves.

#### Language-, country-, gender-, and sexual orientation–based CFAs of the P4 screener

The results are presented in [Supplementary Tables 2 to 5](#), which correspond to language, country, gender, and sexual orientation, respectively. The CFAs supported the structural validity of the P4 screener in 18 of 20 languages, excluding Italian and simplified Mandarin languages. To explore the poor fit in these languages, we conducted subsequent bootstrap exploratory graph analyses (EGAs). We found that in Italian, the dimension stability was high but produced a two-factor solution in 76.02% (factor 1 comprised items 1 and 4; factor 2 comprised items 2 and 3) of the resampling cycles, and a one-factor solution in the remaining 23.98% of cycles; median dimension of 2, standard error = 0.42, CI 1.16, 2.84. In addition, item 2 ‘Have you thought about how you might actually

hurt yourself?’ and item 3 ‘How likely do you think it is that you will act on these thoughts about hurting yourself or ending your life sometime over the next month?’ showed poor stability (0.24), indicating that these items oscillate between 1 and 2 factors. Regarding simplified Mandarin, the bootstrap EGA indicated 100% dimensional and item stabilities, supporting the structural validity of the suicidal ideation scale.

The CFAs supported the structural validity of the suicidal ideation scale in 30 of 31 countries, excluding Italy, as expected based on the language-based CFAs. The CFAs also corroborated the structural validity of the P4 screener in all genders and sexual orientations. Standardized factor loadings by language, country, gender, and sexual orientation are presented in [Supplementary Tables 6–9](#).

### Group differences

Kruskal–Wallis tests indicated significant differences in suicidal ideations between languages,  $\chi^2_{(25)} = 1598.06, P = 1.73^{-322}, \eta^2 = 0.02$  (95% bootstrap CI 0.02, 0.02), countries,  $\chi^2_{(44)} = 1599.16, P = 1.02^{-306}, \eta^2 = 0.02$  (95% bootstrap CI 0.02, 0.02), genders,  $\chi^2_{(4)} = 2185.85, P \sim 0$  (i.e. approximating zero),  $\eta^2 = 0.026$  (95% bootstrap CI 0.02, 0.03), and sexual orientations,  $\chi^2_{(10)} = 4018.17, P \sim 0, \eta^2 = 0.048$  (95% bootstrap CI 0.045, 0.05). Post hoc differences are presented in [Supplementary Tables 10–13](#). [Fig. 1](#) presents the scores for suicidal ideation by country.

#### Reliability and convergent and discriminant validity of the P4 screener

The P4 screener showed adequate reliability in most languages, countries, genders, and sexual orientations ([Supplementary Tables 3–6](#)). The reliability of the scale was less than the recommended level of 0.70 in several languages and in Canada, Mexico, Spain, and Switzerland with regard to countries. However, Coefficient H revealed high reliability in all these instances [Czech (0.98), Hungarian (0.97), Portuguese in Portugal (0.97), and Spanish in Spain (0.98); Canada (0.96), Mexico (0.98), Spain (0.98), and Switzerland (0.96)] indicating that the less-than-adequate alpha and omega scores might stem in several cases from items with lower factor loadings.

The results of the convergence and discriminant validity (i.e. Multi-trait Multimethod analysis) are presented in [Fig. 2](#). The analysis supported high convergence and discriminant validity, such that the average correlation between items of the P4 screener and those of the depression or anxiety scales were moderate in size ( $r_{mean} = 0.35$  for depression,  $r_{mean} = 0.27$  for anxiety), and those assessing substance use were weak in size ( $r_{mean} = 0.07$ ). The average correlation between items of the P4 screener was  $r_{mean} = 0.49$ .

#### Establishing a cutoff score for an ‘At-Risk’ group

The results are summarized in [Fig. 3](#). Bootstrapping the optimal cut-point of the P4 score revealed that using a cut-point of 1 produced a maximum Youden’s index of 0.47, with a sensitivity of 70.45% and a specificity of 76.34% when distinguishing participants ‘at risk’ of a comorbidity of depression and anxiety. Using the cut-point of 1, 2.14 participants were needed to be examined to correctly detect one person with the disorder of interest in a study population of persons with and without the known disorder (i.e. NND value; with a value of 1 as the best possible value). In addition, 4.15 participants needed to be tested for one person to be misdiagnosed by the test (i.e. NNM value). The overall likelihood of being diagnosed compared to being misdiagnosed was 1.94, which indicates adequate effectiveness in the diagnosis process.

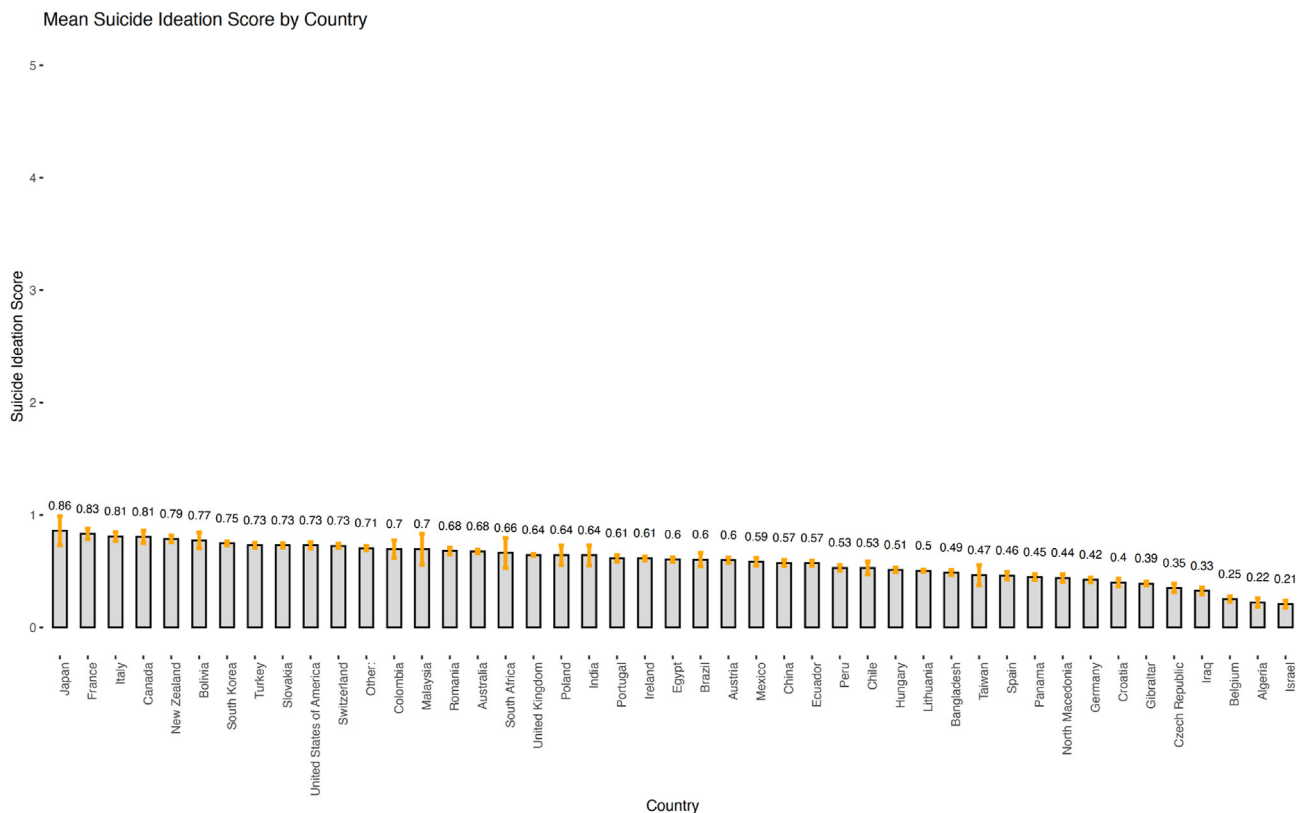


Fig. 1. Suicidal ideation scores by country. Note. Error bars represent standard errors.

**Discussion**

This study validates the P4 suicide screener as a reliable tool for identifying individuals with suicidal ideation across 42 diverse countries. The cross-cultural validation is particularly important, as it allows for the collection of accurate and relevant international data on suicide, which is essential for advancing prevention measures and physician education on suicide prevention.<sup>31</sup> Quality and accessibility of data can impact healthcare professionals’ ability to recognize patterns and identify suicide risk factors, develop effective treatment plans, and make informed decisions for patient care.

Collecting data on suicide mainly from WEIRD countries may not accurately represent the prevalence and underlying factors of suicide in other parts of the world due to variations in suicide rates, risk factors, and cultural and linguistic barriers. Definitions of and discourses on suicide are also largely varying across countries, and not considering these might lead to a biased understanding of the issue. This problem is confounded by the facts that suicide is a sensitive issue and suicidal behavior is illegal in some nations. This study contributes to filling this gap by evaluating the reliability and validity of the P4 suicide screener across 42 diverse countries, including non-WEIRD and developing nations. Using a validated cross-cultural suicide risk assessment is crucial in standardizing and ensuring reliability in identifying high-risk individuals, monitoring changes in risk levels, and ultimately leading to better outcomes for those at risk of suicide. Recognizing the contextual differences between countries and cultures can lead to the development of culturally sensitive suicide prevention strategies tailored to meet the needs of diverse populations, which is crucial for guiding efforts to prevent suicide.<sup>32,33</sup>

Individuals who identified as a sexual minority had the highest frequencies of suicidal ideation, while heterosexual individuals had

the lowest. Findings from the present study also demonstrate that gender-non-conforming individuals exhibited the greatest likelihood of reporting suicidal ideation, followed by individuals who identified as women and men. These results contribute to a growing body of evidence that suggests an increased risk of suicide ideation and behavior among individuals who belong to gender or sexual minorities.<sup>34–39</sup> Healthcare professionals need to be aware of the higher suicide risk among marginalized groups and provide culturally responsive care that addresses their unique needs and experiences. In the present study, we found that 20% reported attempting deliberate self-harm, while 23% reported thinking about specific ways to do so. These observed rates surpass previous prevalence rates,<sup>6</sup> potentially suggesting the inclusion of non-suicidal self-harm thoughts or may derive from the present sample’s diversity (i.e. more sexually diverse individuals were included in the present study compared with previous ones and a heightened rate of suicidal risk was present among sexual minorities).

Finally, the establishment of a clear-cut cutoff score for identifying individuals as ‘at-risk’ for suicidal ideation is a significant aspect of the study, which is essential for early intervention and the provision of appropriate treatment and support in clinical practice. Considering research implications, it is a promising measure to be tested in the clinical field to contribute to overcoming the gap of measures that the predictive ability was tested in clinical studies.<sup>40,41</sup>

*Limitations and future research*

Despite the strengths of the ISS, such as its novelty, large sample size, rigorous methodology, and following open-science practices, some general limitations should be considered in all studies using this data set (for better overview, see [https://osf.io/n3k2c/?view\\_only=838146f6027c4e6bb68371d9d14220b5](https://osf.io/n3k2c/?view_only=838146f6027c4e6bb68371d9d14220b5)). Additional limitations include the reliance on online recruitment, which may

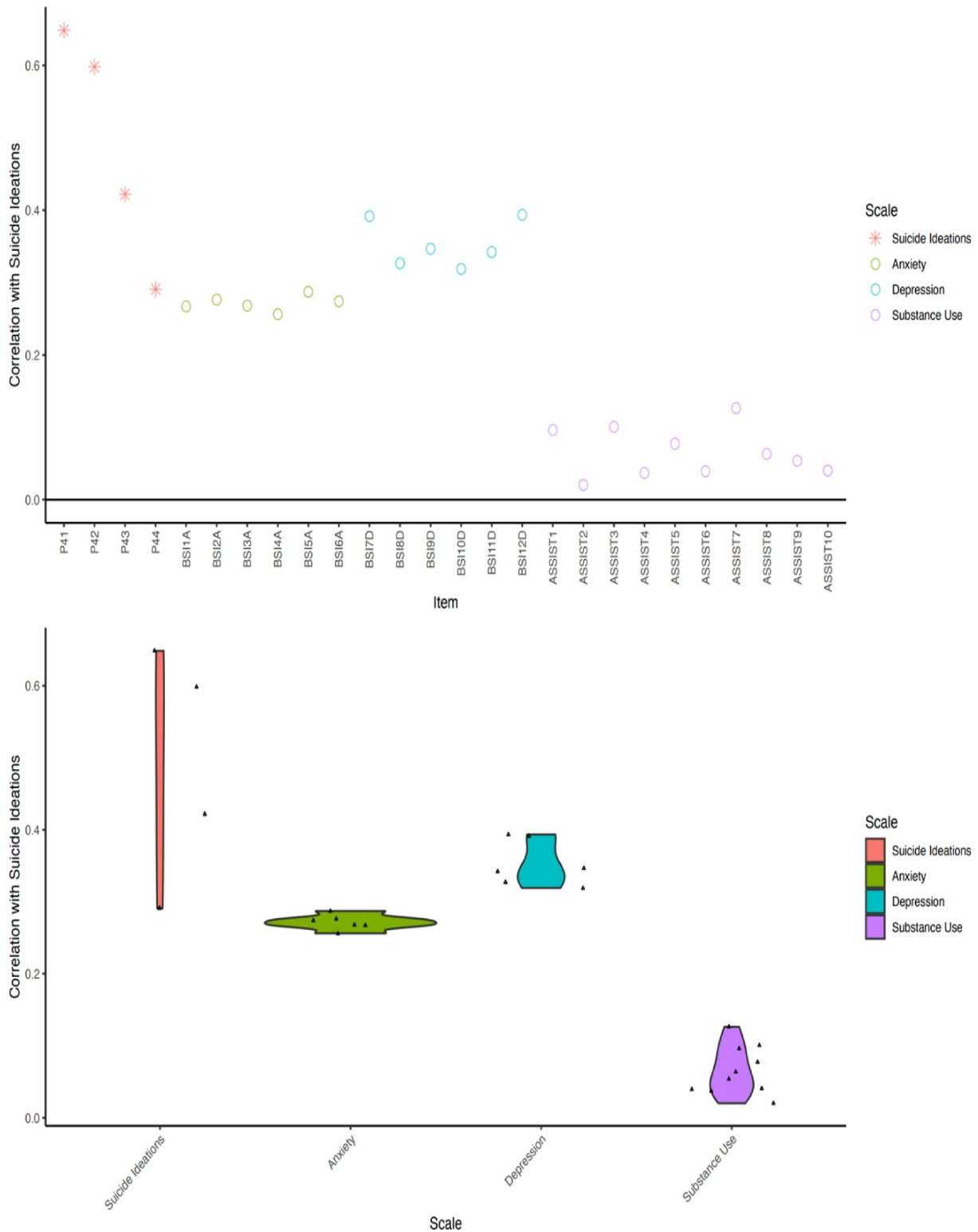
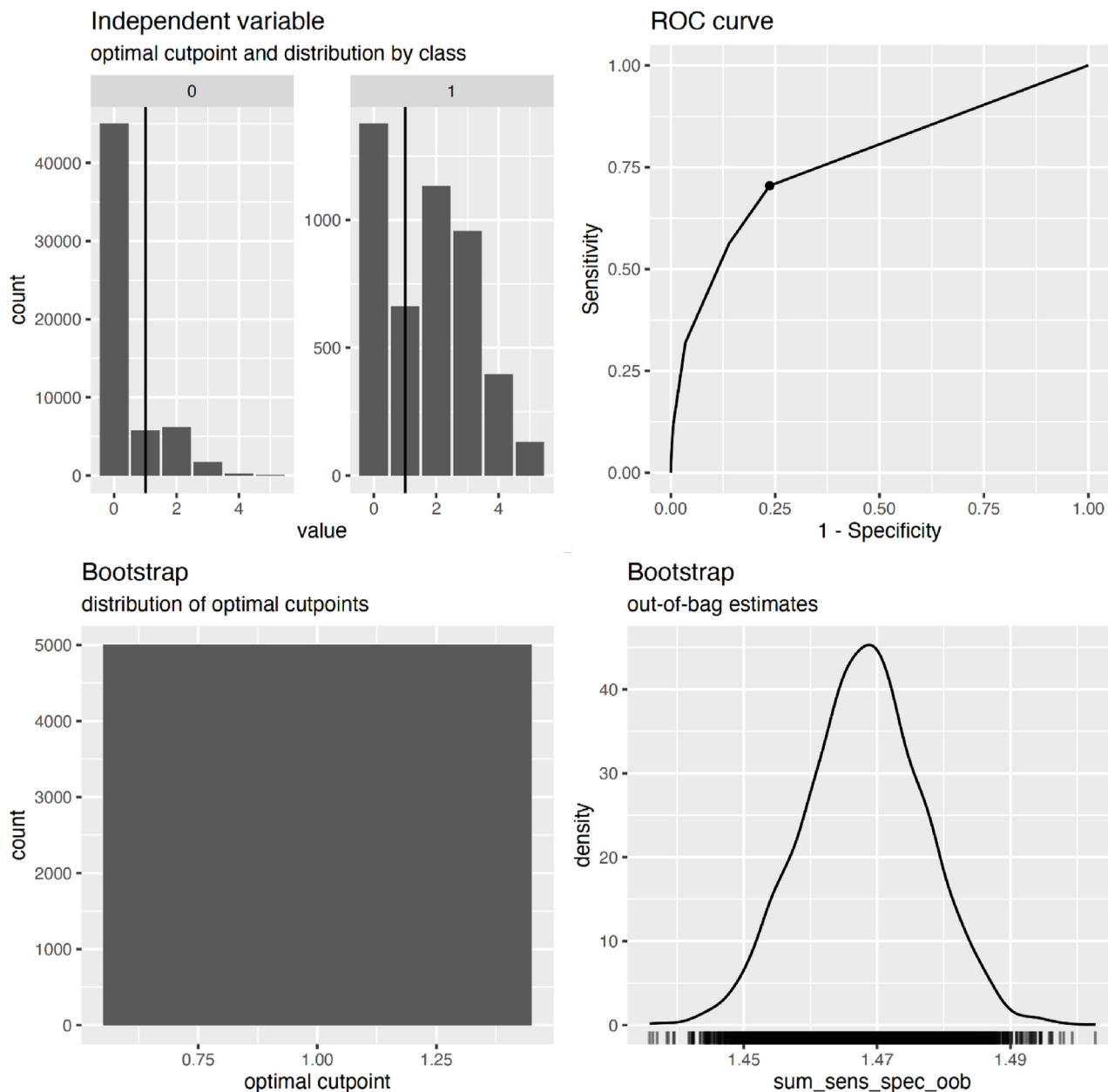


Fig. 2. Multitrait multimethod analysis for assessing convergence and discriminant validity.

introduce selection bias and compromise generalizability, especially to clinical populations. Underrepresentation of specific demographics, such as older adults and those with lower socioeconomic status, is acknowledged, likely due to the online survey approach. Geographical and cultural gaps, notably in Africa and certain Islamic countries, raise concerns about the broader applicability of findings. Furthermore, we did not include a specific

question related to racial minority status in this study. Study-specific limitations include the inability to directly assess suicide, as no established follow-up protocol for evaluating individuals with a high suicide risk was implemented, given the anonymous nature of the study. In addition, the P4 suicide screener does not distinguish between self-harm with and without the intent to die, underscores a limitation, emphasizing the necessity for further





**Fig. 3.** Results of the cut-point optimization process using bootstrap analysis. Top left: density of participants below and above the suggested cut-point of 1 in the scale among participants with “at risk” depression and anxiety (1) and those at low risk (0). Top right: the ROC curve for the estimation process of the optimal cut-point is presented with the black dot indicating the highest Youden’s index. Bottom left: Density of the optimal cut-point in the estimation process. The optimal cut-point was identified as 1 with a standard error of 0, attributable to the discrete nature of the scale. The x-axis spread solely represents a bar over the discrete score of 1. Bottom right: density of the highest summed sensitivity and specificity scores of the suicidal ideation scale.

research to achieve a more precise differentiation. The utilization of more explicit scales may not be suitable in particular cultures, and their use in online anonymous surveys would have raised ethical issues. Addressing these limitations in future research is imperative for a nuanced understanding of suicide ideation rates, particularly in clinical contexts.

**Conclusions**

This study contributes to suicide prevention efforts by demonstrating the validity and reliability of the P4 suicide screener, an efficient tool for identifying individuals at risk for suicidal ideation and behavior across diverse populations. The use of a standardized tool may improve the quality and accessibility of data on suicide,

which is important for developing and improving the effectiveness of suicide prevention strategies in clinical practice.

**Notes**

All published papers and conference presentations using the International Sex Survey data set are available for public viewing on the related OSF pages.

- 1) Publications: [https://osf.io/jb6ey/?view\\_only=0014d87bb2b546f7a2693543389b934d](https://osf.io/jb6ey/?view_only=0014d87bb2b546f7a2693543389b934d);
- 2) Conference presentations: [https://osf.io/c695n/?view\\_only=7cae32e642b54d049e600ceb8971053e](https://osf.io/c695n/?view_only=7cae32e642b54d049e600ceb8971053e).

## Author statements

### Acknowledgments

The authors would like to thank Anastasia Lucic for her help with project administration and data collection and Natasha Zippan, Yasmin Benoit, Raquel Savage, Anne-Marie Menard, Club Sexu, Telex. hu, and La Voix du Nord, L'Avenir de l'Artois, Wéo TV, and Le Parisien for their help with data collection.

### Ethical approval

The authors assert that all procedures contributing to this work comply with the relevant national and institutional committees' ethical standards on human experimentation and the Helsinki Declaration. The study was approved by all collaborating countries' national/institutional ethics review boards: [https://osf.io/n3k2c/?view\\_only=838146f6027c4e6bb68371d9d14220b5](https://osf.io/n3k2c/?view_only=838146f6027c4e6bb68371d9d14220b5).

### Funding

C.-Y.L. was supported by the WUN Research Development Fund (RDF) 2021 and the Higher Education Sprout Project, the Ministry of Education at the Headquarters of University Advancement at the National Cheng Kung University (NCKU); J.B. was supported by the WUN 2021 grant. C.L. was supported by the WUN Research Development Fund (RDF) 2021; G.O. was supported by the ANR grant of the Chaire Professeur Junior of Artois University and by the Strategic Dialogue and Management Scholarship (Phase 1 and 2); G.C.Q.G. was supported by the SNI #073 -2022 (SENACYT, Rep. of Panama); H.F. was supported by Grant-in-Aid for Transformative Research Areas (A) (Japan Society for The Promotion of Science, JP21H05173), Grant-in-Aid for Scientific Research (B) (Japan Society for The Promotion of Science, 21H02849), and the smoking research foundation; J.B.G. was supported by grants from the Kindbridge Research Institute, the International Center for Responsible Gaming, and the Problem Gambling Network of Ohio; K. Lu was supported by the Charles University institutional support programme Cooperatio-Health Sciences; K. Le was supported by the Sonatina grant awarded by National Science Centre, Poland, grant number: 2020/36/C/HS6/00005.; K.R. was supported by funding from the Hauts-de-France region (France) called "Dialogue Stratégique de Gestion 2 (DSG2)"; L.C. was supported by the National Social Science Foundation of China (Grant No. 19BSH117 ); L.N. was supported by the ÚNKP-22-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.; M.G. was supported by National Science Centre of Poland grant nb. 2021/40/Q/HS6/00219; M.K. was supported by the ÚNKP-22-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.; P.M.-T. was supported by Universidad Científica del Sur; R.C. was supported by RC was supported by Auckland University of technology, 2021 Faculty Research Development Fund; R.G. was supported by Charles University institutional support programme Cooperatio-Health Sciences; S.A.D. was supported by the Tempus Public Foundation; S.B. was supported by a Tier 1 Canada Research Chair; S.U.s.r.t. was supported by Brain Korea 21 (BK21) program of National Research Foundation of Korea; S.W.K. was supported by the Kindbridge Research Institute.; Z.D. was supported by the Hungarian National Research, Development, and Innovation Office (Grant numbers: KKP126835).

### Competing interests

The authors declare no conflict of interest with the content of this manuscript. S.W.K. discloses that he has received funding from the International Center for Responsible Gaming, MGM Resorts International, Center for the Application of Substance Abuse Technologies, Taylor Francis, Springer Nature, the Nevada Problem Gambling Project, Sports Betting Alliance, and Kindbridge Research Institute. Dr. Potenza discloses that he has consulted for and advised Game Day Data, Addiction Policy Forum, AXA, Idorsia, Baria-Tek, and Opiant Therapeutics; been involved in a patent application involving Novartis and Yale; received research support from the Mohegan Sun Casino, Children and Screens and the Connecticut Council on Problem Gambling; consulted for or advised legal and gambling entities on issues related to impulse control, internet use and addictive behaviors; performed grant reviews; edited journals/journal sections; given academic lectures in grand rounds, CME events, and other clinical/scientific venues; the University of Gibraltar receives funding from the Gibraltar Gambling Care Foundation, an independent not-for-profit charity generated books or chapters for publishers of mental health texts. The University of Gibraltar receives funding from the Gibraltar Gambling Care Foundation, an independent not-for-profit charity. However, these funding sources are not related to this study, and the funding institution had no role in the study design or the collection, analysis, and interpretation of the data, the writing of the manuscript, or the decision to submit the paper for publication.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.puhe.2023.12.031>.

### References

- World Health Organization. *Suicide*. <https://www.who.int/news-room/fact-sheets/detail/suicide>, 2021. [Accessed 4 January 2023].
- Favril L, Yu R, Uyar A, Sharpe M, Fazel S. Risk factors for suicide in adults: systematic review and meta-analysis of psychological autopsy studies. *Evid Base Ment Health* 2022;10:1–8. <https://doi.org/10.1136/ebmental-2022-300549>.
- Too LS, Spittal MJ, Bugeja L, Reifels L, Butterworth P, Pirkis J. The association between mental disorders and suicide: a systematic review and meta-analysis of record linkage studies. *J Affect Disord* 2019;259:302–13. <https://doi.org/10.1016/j.jad.2019.08.054>.
- Gili M, Castellví P, Vives M, de la Torre-Luque A, Almenara J, Blasco MJ, et al. Mental disorders as risk factors for suicidal behavior in young people: a meta-analysis and systematic review of longitudinal studies. *J Affect Disord* 2018. <https://doi.org/10.1016/j.jad.2018.10.115>.
- Holman MS, Williams MN. Suicide risk and protective factors: a network approach. *Arch Suicide Res* 2020;1–18. <https://doi.org/10.1080/13811118.2020.1774454>.
- Bullington C, Kroenke K. P4 suicidality screener: literature synthesis and results from two randomized trials. *Gen Hosp Psychiatr* 2023;85:177–84. <https://doi.org/10.1016/j.genhosppsych.2023.11.004>.
- Dube P, Kroenke K, Bair MJ, Theobald D, Williams LS. The P4 screener: evaluation of a brief measure for assessing potential suicide risk in 2 randomized effectiveness trials of primary care and oncology patients. *Prim Care Companion J Clin Psychiatry* 2010;12(6). <https://doi.org/10.4088/PCC.10M00978BLU>.
- Chu JP, Goldblum P, Floyd R, Bongar B. The cultural theory and model of suicide. *Appl Prev Psychol* 2010;14(1–4):25–40. <https://doi.org/10.1016/j.appsy.2011.11.001>.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25(24):3186–91.
- Böthe B, Koós M, Nagy L, Kraus SW, Potenza MN, Demetrovics Z. International Sex Survey: study protocol of a large, cross-cultural collaborative study in 45 countries. *J Behav Addict* 2022;10:632–45. <https://doi.org/10.1556/2006.2021.00063>.
- Böthe B, Koós M, Nagy L, Kraus SW, Potenza MN, Demetrovics Z. International Sex Survey: study protocol of a large, cross-cultural collaborative study in 45 countries. *J Behav Addict* 2021;10(3):632–45. <https://doi.org/10.1556/2006.2021.00063>.
- Asner-Self KK, Schreiber JB, Marotta SA. A cross-cultural analysis of the brief symptom inventory-18. *Cult Divers Ethnic Minor Psychol* 2006;12(2):367–75. <https://doi.org/10.1037/1099-9809.12.2.367>.

13. Group, W.A.W. The Alcohol, smoking and substance involvement screening test (ASSIST): development, reliability and feasibility. *Addiction* 2002;**97**(9): 1183–94. <https://doi.org/10.1046/j.1360-0443.2002.00185.x>.
14. Lee CT, Lin CY, Koós M, Nagy L, Kraus SW, Demetrovics Z, et al. The eleven-item Alcohol, smoking and substance involvement screening test (ASSIST-11): cross-cultural psychometric evaluation across 42 countries. *J Psychiatr Res* 2023;**165**:16–27. <https://doi.org/10.1016/j.jpsychires.2023.06.033>.
15. Little RJ. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc* 1988;**83**(404):1198–202.
16. Newman DA. Missing data: five practical guidelines. *Organ Res Methods* 2014;**17**(4):372–411. <https://doi.org/10.1177/1094428114548590>.
17. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Testing structural equation models* 1993;**21**(2):136–62. <https://doi.org/10.1167/jovs.04-1279>.
18. Marsh HW, Hau KT, Grayson D. Goodness of fit evaluation in structural equation modeling. In: Maydeu-Olivares A, McArdle J, editors. *Contemporary psychometrics*. Erlbaum; 2005. p. 275–340.
19. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res* 2003;**8**(2):23–74.
20. Sahoo M. Structural equation modeling: threshold criteria for assessing model fit. In: *Methodological issues in management research: advances, challenges, and the way ahead*. Emerald Publishing Limited; 2019.
21. Kiliç A, Uysal İ, Atar B. Comparison of confirmatory factor analysis estimation methods on binary data. *International Journal of Assessment Tools in Education* 2020: 451–87. <https://doi.org/10.21449/ijate.660353>. Published online August 24.
22. Golino H, Shi D, Christensen AP, Garrido LE, Nieto MD, Sadana R, et al. Investigating the performance of Exploratory Graph Analysis and traditional techniques to identify the number of latent factors: a simulation and tutorial. *Psychol Methods* 2019;**25**(3):292–320. <https://doi.org/10.1037/met0000255>.
23. Golino HCAP MR. EGAnet: exploratory graph analysis: a framework for estimating the number of dimensions in multivariate data using network psychometrics. *R package version* 2020;**9**(5). Published online.
24. Revelle W. CDM. Supplemental material for reliability from  $\alpha$  to  $\omega$ : a tutorial. *Psychol Assess* 2019. <https://doi.org/10.1037/pas0000754> [suppl].
25. Kalkbrenner MT. Alpha, omega, and H internal consistency reliability estimates: reviewing these options and when to use them. *Counsel Outcome Res Evaluat* 2021. <https://doi.org/10.1080/21501378.2021.1940118>.
26. Falissard B. Psy: various procedures used in psychometrics. In: *Version R package Version 1.2.*; 2022.
27. Cohen J. *A Power Primer***112**; 1992.
28. Linn S, Grunau PD. New patient-oriented summary measure of net total gain in certainty for dichotomous diagnostic tests. *Epidemiol Perspect Innovat* 2006;**3**. <https://doi.org/10.1186/1742-5573-3-11>.
29. Habibzadeh F, Yadollahie M. Number needed to misdiagnose: a measure of diagnostic test effectiveness. *Epidemiology* 2013;**24**(1):170. <https://doi.org/10.1097/EDE.0b013e31827825f2>.
30. Citrome L, Ketter TA. When does a difference make a difference? Interpretation of number needed to treat, number needed to harm, and likelihood to be helped or harmed. *Int J Clin Pract* 2013;**67**(5):407–11. <https://doi.org/10.1111/ijcp.12142>.
31. Mann JJ, Apter A, Bertolote J, Beautrais A, Currier D, Haas A, et al. Suicide prevention strategies A systematic review. *JAMA* 2005;**294**(16):2064–74. <https://jamanetwork.com/>.
32. Hjelmeland H. Cultural context is crucial in suicide research and prevention. *Crisis* 2011;**32**(2):61–4. <https://doi.org/10.1027/0227-5910/a000097>.
33. Dückers MLA, Reifels L, De Beurs DP, Brewin CR. The vulnerability paradox in global mental health and its applicability to suicide. *Br J Psychiatry* 2019;**215**(4):588–93. <https://doi.org/10.1192/bjp.2019.41>.
34. Ramchand R, Schuler MS, Schoenbaum M, Colpe L, Ayer L. Suicidality among sexual minority adults: gender, age, and race/ethnicity differences. *Am J Prev Med* 2022;**62**(2):193–202. <https://doi.org/10.1016/j.amepre.2021.07.012>.
35. Mcdaniel JS, Purcell D, D'augelli AR. The relationship between sexual orientation and risk for suicide: research findings and future directions for research and prevention. *Suicide Life-Threatening Behav* 2001:84–105.
36. Plöderl M, Wagenmakers EJ, Tremblay P, Ramsay R, Kralovec K, Fartacek C, et al. Suicide risk and sexual orientation: a critical review. *Arch Sex Behav* 2013;**42**:715–27. <https://doi.org/10.1007/s10508-012-0056-y>.
37. Salway T, Plöderl M, Liu J, Gustafson P. Effects of multiple forms of information bias on estimated prevalence of suicide attempts according to sexual orientation: an application of a bayesian misclassification correction method to data from a systematic review. *Am J Epidemiol* 2019;**188**(1):239–49. <https://doi.org/10.1093/aje/kwy200>.
38. McNeil J, Ellis SJ, Eccles FJR. Suicide in trans populations: a systematic review of prevalence and correlates. *Psychol Sex Orientat Gend Divers* 2017;**4**(3):341–53. <https://doi.org/10.1037/sgd0000235>.
39. Turban JL, Beckwith N, Reisner SL, Keuroghlian AS. Association between recalled exposure to gender identity conversion efforts and psychological distress and suicide attempts among transgender adults. *JAMA Psychiatr* 2020;**77**(1):68–76. <https://doi.org/10.1001/jamapsychiatry.2019.2285>.
40. Bolton JM, Gunnell D, Turecki G. Suicide risk assessment and intervention in people with mental illness. *BMJ* 2015. <https://doi.org/10.2307/26524477>.
41. Randall JR, Colman I, Rowe BH. A systematic review of psychometric assessment of self-harm risk in the emergency department. *J Affect Disord* 2011;**134**(1–3):348–55. <https://doi.org/10.1016/j.jad.2011.05.032>.
42. Quintana GR, Ponce FP, Escudero-Pastén JI, Santibáñez-Palma JF, Nagy L, Koós M, et al. Cross-cultural validation and measurement invariance of anxiety and depression symptoms: a study of the Brief Symptom Inventory (BSI) in 42 countries. *J Affect Disorders* 2024.
43. Horváth Z, Nagy L, Koós M, Kraus SW, Demetrovics Z, Potenza MN, et al. Psychometric properties of the Alcohol Use Disorders Identification Test (AUDIT) across cross-cultural subgroups, genders, and sexual orientations: findings from the International Sex Survey (ISS). *Compr Psychiatr* 2023;**127**:152427.