

## ORIGINAL ARTICLE

# Falling between the cracks: The effect of using different levels of suicide risk exclusion criteria on sample characteristics when recruiting for an online intervention for depression

Alexandra Godinho MScCH<sup>1</sup> | Christina Schell BSc<sup>1</sup> | John A. Cunningham PhD<sup>1,2,3</sup> 

<sup>1</sup>Institute of Mental Health and Policy Research, Centre for Addiction and Mental Health, Toronto, ON, Canada

<sup>2</sup>National Addiction Centre, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, United Kingdom

<sup>3</sup>Department of Psychiatry, University of Toronto, Toronto, ON, Canada

## Correspondence

John A. Cunningham, Institute of Psychiatry, Psychology & Neuroscience, 16 De Crespigny Park, London, SE5 8AF  
Email: John.Cunningham@kcl.ac.uk

## Funding information

This research is funded by the Canadian Institutes of Health Research grant number PJT 153 324; however, the findings and conclusions of this paper are solely of the authors and do not necessarily represent the views of the Canadian Institutes of Health Research. The funding body has no role or influence on the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Abstract

**Background:** Despite a strong link between suicide risk and depression, a recent literature review found that many effectiveness studies for online depression interventions exclude individuals at risk of suicide. This study scrutinizes how different suicide risk exclusion criteria impact recruitment rates and final sample characteristics.

**Materials and Methods:** Two recruitment periods for an online depression intervention trial utilized different suicide risk cutoff exclusion criteria, a one-point difference on the last item of the Personal Health Questionnaire (i.e., more than 0 (Not at all) vs. more than 1 (Several Days)). Bivariate statistics were used to assess differences in recruitment rates and sample characteristics between these two recruitment periods, while all other eligibility criteria and recruitment strategies remained consistent.

**Results:** The recruitment period using the least restrictive suicide risk exclusion criteria yielded twice as many participants; however, recruited sample characteristics did not significantly differ among demographic or clinical characteristics, despite observable trends.

**Discussion:** Researchers should carefully select suicide risk exclusion criteria that balance recruitment rates, study budgets, and sample selection biases, while minimizing participant harm. Moreover, researchers are urged to report suicide risk exclusion rates and consider these exclusions when interpreting results. Limitations of the results are also discussed.

## INTRODUCTION

It has generally become accepted that Internet interventions for depression and other mental disorders can be effective, especially when some form of human guidance is provided. A literature review of mobile apps and web-based therapies for depression concluded that comparable treatments appear to show little difference in efficacy

when delivered via different formats (e.g., online vs. face-to-face) and furthermore that completely self-guided online interventions show small yet significant positive effects (Cuijpers et al., 2017). While these findings are promising, much of the existing evidence relies on web-based intervention studies that have overlooked an important subgroup of depressed individuals, those who may be at risk for suicide.

Trial Registration: ClinicalTrials.gov NCT03421080; Participant recruitment started April 26, 2018.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Suicide and Life-Threatening Behavior* published by Wiley Periodicals LLC on behalf of American Association of Suicidology

In a recent review of randomized controlled trials (RCTs) of online interventions for depression, Sander et al. (2020) found that the majority of studies, over 70%, excluded individuals at risk of suicide. Furthermore, an independent review of these studies by the authors revealed that less than 60% of researchers who exclude those at risk of suicide report the number of excluded individuals; and when reported, suicide risk represented a substantial amount of excluded screeners (between 5% and 51%) and up to one third of all eligible participants.

While researchers may have important reasons for excluding these individuals, such as a lack of resources for supporting them (e.g., time, expertise, expense) and resistance from ethics committees, it is unknown how their exclusion may confound study findings or alter sample characteristics. Taking into account that depressed individuals may be at a 25 times greater risk of suicide (American Association of Suicidology, 2014) and that suicidal ideation has been found to be a predictor of treatment seeking for major depression (Magaard et al., 2017), excluding this potentially large subgroup can both limit the external validity of the findings and artificially inflate or mute the interventions' efficacy (Andriessen et al., 2019; Lakeman & Fitzgerald, 2009; Sisti & Joffe, 2018; Zimmerman et al., 2005). Consequently, it is imperative that researchers better understand how the exclusion of individuals at risk of suicide impacts the sample characteristics and recruitment rate of individuals into online depression interventions, which is what this paper aims to address. In an ongoing study testing the effectiveness of an online intervention for depression with and without the provision of a brief alcohol treatment, the authors utilized the last item of the Personal Health Questionnaire (PHQ) to determine suicide risk. During the first four months of recruitment, participants scoring greater than zero (Not at all) on the last item were excluded; however, due to difficulties recruiting, this criterion was expanded to exclude only those scoring greater than one (Several days). As all other parts of the recruitment process and eligibility criteria were identical, it provided a unique opportunity to examine how a small change in suicide risk criteria impacts recruitment rates and subsequently sample characteristics. Due to the exploratory nature of this paper, no a priori hypotheses were made.

## MATERIALS AND METHODS

### Study design

The sample analyzed in this paper is part of a larger ongoing research study for which a study protocol has been published elsewhere (Cunningham et al., 2018). Briefly, individuals

currently feeling persistent low mood or depression were invited to participate in a study to “help improve an online intervention for depression” via online advertisements which commenced in April 2019. Eligibility criteria for the study consisted of being 18 years or older, scoring 10 or more on the PHQ (Kroenke et al., 2001), scoring 8 or more on the Alcohol Use Disorders Identification Test (AUDIT) (Saunders et al., 1993), and agreeing to be followed up for 6 months. Additionally, potential participants were screened for suicidal ideation using the last item of the PHQ-9 “Over the last 2 weeks, how often have you been bothered by thoughts that you would be better off dead or of hurting yourself in some way,” with the response options Not at all (0), Several days (1), More than half the days (2), and Nearly every day (3). For those recruited between April and July 2019, exclusion criteria included scoring greater than zero on this item, while those recruited after August 2019 were only excluded if they scored more than one.

Those found eligible were also required to provide the study team with contact information (e.g., email, mailing address, telephone number) for verification purposes and complete an online consent form. Once consent was provided, eligible participants were then asked to complete a baseline questionnaire measuring both clinical (i.e., depression, anxiety, alexithymia, past treatment use, quality of life) and demographic characteristics and were randomized to receive an online intervention for depression (i.e., MoodGYM) with or without the additional provision of a brief intervention for alcohol use. Only those who accessed the online intervention(s) were officially enrolled into the study and compensated with a \$10 Amazon.ca gift card. In addition, participants who completed the three- and six-month follow-up surveys were compensated with a \$20 and \$30 Amazon.ca gift card, respectively. It is important to note that individuals who were not found eligible to participate in the study were provided free access to the online depression intervention (i.e., MoodGYM) for the duration of the study. Ethics approval for the research methods used in this study was provided by the standing ethics review committee of the Centre for Addiction and Mental Health (CAMH).

### Data analysis

As advertising campaigns, including monthly budgets and ad content, were comparable throughout the sampling, individuals who were screened for inclusion into the study during the first four months of recruitment (April to July 2019) and those screened during the subsequent four months (August to October 2019) were compared using rates of enrollment. A series of chi-square tests were conducted to determine whether the proportion of eligible screeners, completed consents, and intervention registrations differed between the

two groups. In addition, those who were recruited into the study from each recruitment period were compared on demographic and clinical characteristics using independent *t* tests and chi-square tests. All analyses were performed using SPSS, version 25.0.

## RESULTS

A total of 5341 screening surveys were completed between April 2019 and October 2019, of which 478 were found eligible to participate, and 316 went on to provide contact information. Overall, 220 potential participants provided consent, and 170 signed up for the intervention and were officially enrolled into the study. The proportion of individuals which were excluded at different stages of recruitment was compared between the first and subsequent four months in Table 1. Overall, a smaller proportion of individuals were found eligible during the first four months of recruitment as compared to the subsequent four months  $\chi^2(1, N = 5341) = 142.03, p < 0.001$ . In particular, individuals completing the screener during the first four months were more likely to be screened out for being less than 18 years of age ( $p < 0.001$ ), experiencing suicidal ideation ( $p < 0.001$ ), and having a score less than 8 on the AUDIT ( $p < 0.001$ ), as compared to those screened within the subsequent four months. However, post-eligibility, differences observed between the two recruitment periods did not carry forward to the proportion of those who subsequently completed a consent form ( $p = 0.677$ ), or accessed the online intervention and were officially enrolled into the study ( $p = 0.631$ ). In addition, the recruited samples from the different recruitment periods did not display any significant differences in their demographic or clinical characteristics (Table 2).

## DISCUSSION

The majority of research trials of online interventions for depression exclude individuals at risk of suicide (Sander et al., 2020); however, this paper sought to determine how different suicide risk exclusion criteria may impact recruitment rates and overall sample characteristics. In general, we found that less restrictive suicide risk exclusion criteria more than doubled the proportion of eligible participants; however, this did not appear to have any effect on the proportion of individuals who completed subsequent stages of recruitment such as providing informed consent or accessing the study intervention. Moreover, no significant differences were observed between the two recruited groups among other mental health symptoms (e.g., depression, anxiety), alcohol use, past treatment use, or demographic characteristics. Indeed, it was quite surprising that the two recruitment groups did not differ on depression symptoms as measured by the PHQ-8 despite less restrictions being placed on only one of the groups. It would have been expected that allowing participants to have some suicidal ideation would have increased that recruited group overall depression symptoms as arguably it should have been a more severely depressed sample. The expected trend, although not significant, was observed within the anxiety scores and quality of life between the groups, as loosening the PHQ criteria lead to recruiting a group that reported slightly greater anxiety and a lower quality of life. As this study was somewhat underpowered to significantly detect these differences, future research should address this in more detail and disentangle the relationship between loosening suicidal ideation on the PHQ and its impact on recruited sample characteristics.

While these findings alone cannot recommend what suicide risk level is optimal for use as exclusion criteria, or

<i>n</i> (%)	Months 1–4 of recruitment ( <i>N</i> = 2914)	Months 5–8 of recruitment ( <i>N</i> = 2427)	<i>p</i>
Total # of surveys ineligible <sup>a</sup>			
Age <18	207 (7.1)	89 (3.7)	<0.001
PHQ <10	356 (13.2)	303 (13.0)	0.841
PHQ item #9 <sup>b</sup>	1620 (68.9)	750 (36.9)	<0.001
AUDIT <8	594 (81.3)	944 (73.5)	<0.001
Total # eligible participants	137 (4.7)	341 (14.1)	<0.001
Provided consent	61 (44.5)	159 (46.6)	0.677
Officially enrolled <sup>c</sup>	51 (37.2)	119 (34.9)	0.631

**TABLE 1** Comparison of recruitment rates across between first four months of recruitment and subsequent four months.

Abbreviations: AUDIT, Alcohol Use Disorders Identification Test; PHQ, Personal Health Questionnaire.

<sup>a</sup>Individuals were screening out of study in the order ineligible criteria are presented and were not required to complete the remainder of the screening questionnaire. Therefore, proportions reflect % of those asked the question and not all screeners.

<sup>b</sup>Assessed as =0 for months 1–4 of recruitment, and as <2 for months 5–8 of recruitment.

<sup>c</sup>Accessed intervention.

**TABLE 2** Sample characteristic differences between participants recruited during first four months vs. subsequent four months.

	Months 1–4 of recruitment ( <i>N</i> = 51)	Months 5–8 of recruitment ( <i>N</i> = 119)	<i>p</i>
Age, Mean ( <i>SD</i> )	40.2 (13.2)	39.0 (13.8)	0.601
Male, % ( <i>n</i> )	21.6 (11)	33.6 (40)	0.116
Post-secondary education, % ( <i>n</i> )	76.5 (39)	68.1 (81)	0.270
Married/Common-law, % ( <i>n</i> )	43.1 (22)	30.3 (36)	0.104
Full/Part-time employed, % ( <i>n</i> )	64.7 (33)	60.5 (72)	0.605
Family income less than \$30,000, % ( <i>n</i> )	38.0 (19)	43.9 (50)	0.484
Past depression treatment, % ( <i>n</i> )	70.6 (36)	67.2 (80)	0.666
Past alcohol use treatment, % ( <i>n</i> )	21.6 (11)	25.2 (30)	0.611
AUDIT <sup>a</sup> , mean ( <i>SD</i> )	13.4 (5.4)	14.0 (5.6)	0.517
CES-D <sup>b</sup> , mean ( <i>SD</i> )	31.2 (7.5)	31.5 (7.3)	0.794
GAD <sup>c</sup> , mean ( <i>SD</i> )	11.4 (5.1)	12.6 (4.9)	0.154
PHQ-8 <sup>d</sup>	15.8 (3.9)	15.9 (4.0)	0.875
EUROHIS-QoL <sup>e</sup> , mean ( <i>SD</i> )	22.3 (5.3)	20.9 (4.8)	0.098
Toronto Alexithymia Scale, mean ( <i>SD</i> )	57.4 (12.0)	56.9 (10.9)	0.785

<sup>a</sup>Alcohol Use Disorders Identification Test.

<sup>b</sup>Center for Epidemiologic Studies Depression Scale (CES-D), NIMH.

<sup>c</sup>Generalized Anxiety Disorder 7-item scale.

<sup>d</sup>Due to differences in recruitment cutoff scores, the PHQ-8 was compared across the groups.

<sup>e</sup>European Health Interview Survey—Quality of Life 8-item; higher scores indicate a better quality of life.

whether individuals should be excluded at all, it provides a learning opportunity for researchers. Potentially excluding nearly half of all eligible participants as a result of a one-point difference on a 4-point suicide risk item can have important implications on the validity of the data collected, as well as the results of the research, as it can introduce sample selection biases. Moreover, the exclusion of such a large proportion of individuals can also compromise study budgets and timelines as recruiting individuals for online depression interventions can be expensive, time-consuming, and challenging, particularly when subsamples of depressed individuals are being recruited as was the case in this study (i.e., those with hazardous drinking). Therefore, we recommend that researchers carefully balance the use of less restrictive suicide risk criteria, while ensuring that the necessary resources are in place for those who may be at risk of suicide. It is important that researchers balance recruitment goals and data validity, all while minimizing potential harm to participants. In addition, we strongly encourage researchers to be transparent about how many individuals are excluded from their research for being at risk of suicide, how this risk was measured, and

consider possible sample selection biases when interpreting results.

While this study provided a unique opportunity to examine how different suicide risk cutoff scores impact recruitment rates and sample characteristics, some important limitations should be noted. Firstly, some have questioned the use of the last item of the PHQ-9 for determining suicidal risk in recent literature (Na et al., 2018). However, while it continues to be one of the most popular methods of assessing suicidal risk, we recommend that this research also be replicated using other suicide risk measures that can be utilized online such as item 9 on the Beck Depression Inventory, or suicide questionnaires (see (Beck et al., 1961, 1988; Gega et al., 2005). Secondly, eligibility criteria for this study included hazardous drinking; therefore, results are not generalizable to other depression samples recruited online. Thirdly, despite efforts to maintain similar recruitment strategies throughout the project, it is unclear why significantly more individuals were screened out for being <18 during the first four months of recruitment and whether this impacted other observed outcomes.

## CONCLUSION

Given the prevalence of suicide risk among depression samples and its possible role in treatment seeking, researchers are urged to use less restrictive suicide risk exclusion criteria. This improves recruitment rates, and thus study budgets and timelines, and may also minimize the risk of sample selection biases while ensuring that individuals at serious risk of suicide are still excluded. Future research should examine the impact of different suicide risk exclusion criteria on ongoing participant engagement with online depression interventions as well as treatment outcomes.

## ACKNOWLEDGMENTS

Support to CAMH for salary and infrastructure has been provided by the Ministry of Health and Long-Term Care.

## CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

## ETHICS APPROVAL

Ethics approval for the research methods used in this study was provided by the standing ethics review committee of the Centre for Addiction and Mental Health (CAMH) (REB Protocol #064-2017).

## ORCID

John A. Cunningham  <https://orcid.org/0000-0002-0668-5982>

## REFERENCES

- American Association of Suicidology (Producer). (2014). Depression and Suicide Risk.
- Andriessen, K., Reifels, L., Kryszynska, K., Robinson, J., Dempster, G., & Pirkis, J. (2019). Ethical concerns in suicide research: Results of an international researcher survey. *Empirical Research on Research Ethics, 14*, 383–394.
- Beck, A. T., Steer, R. A., & Ranieri, W. F. (1988). Scale for suicide ideation: psychometric properties of a self-report version. *Journal of Clinical Psychology, 44*, 499–505.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561–571.
- Cuijpers, P., Kleiboer, A., Karyotakim, E., & Reper, H. (2017). Internet and mobile interventions for depression: Opportunities and challenges. *Depression and Anxiety, 34*, 596–602.
- Cunningham, J. A., Hendershot, C. S., Kay-Lambkin, F., Neighbors, C., Griffiths, K. M., Bennett, K., Bennett, A., Godinho, A., & Schell, C. (2018). Does providing a brief internet intervention for hazardous alcohol use to people seeking online help for depression reduce both alcohol use and depression symptoms among participants with these co-occurring disorders? Study protocol for a randomised controlled trial. *British Medical Journal Open, 8*, e022412.
- Gega, L., Kenwright, M., Mataix-Cols, D., Cameron, R., & Marks, I. M. (2005). Screening people with anxiety/depression for suitability for guided self-help. *Cognitive Behaviour Therapy, 34*, 16–21.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606–613.
- Lakeman, R., & Fitzgerald, M. (2009). Ethical suicide research: A survey of researchers. *International Journal of Mental Health Nursing, 18*, 10–17.
- Magaard, J. K., Seeralan, T., Schulz, H., & Brutt, A. L. (2017). Factors associated with help-seeking behaviour among individuals with major depression: A systematic review. *PLoS One, 12*, e0176730.
- Na, P. J., Yaramala, S. R., Kim, J. A., Kim, H., Goes, F. S., Zandi, P. P., Vande Voort, J. L., Sutor, B., Croarkin, P., & Bobo, W. V. (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). *Journal of Affective Disorders, 232*, 34–40.
- Sander, L., Gerhardinger, K., Bailey, E., Robinson, J., Lin, J., Cuijpers, P., & Mühlmann, C. (2020). Suicide risk management in research on internet-based interventions for depression: A synthesis of the current state and recommendations for future research. *Journal of Affective Disorders, 263*, 676–683. <https://doi.org/10.1016/j.jad.2019.11.045>
- Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on early detection of persons with Harmful Alcohol Consumption—II. *Addiction, 88*, 791–804.
- Sisti, D. A., & Joffe, S. (2018). Implications of zero suicide for suicide prevention research. *JAMA, 320*(16), 1633–1634.
- Zimmerman, M., Chelminski, I., & Posternak, M. A. (2005). Generalizability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. *The American Journal of Psychiatry, 162*(7), 1370–1372.

**How to cite this article:** Godinho A, Schell C, Cunningham JA. Falling between the cracks: The effect of using different levels of suicide risk exclusion criteria on sample characteristics when recruiting for an online intervention for depression. *Suicide Life Threat Behav.* 2021;51:736–740. <https://doi.org/10.1111/sltb.12761>