Journal of Affective Disorders Reports 16 (2024) 100763

Contents lists available at ScienceDirect

Journal of Affective Disorders Reports

journal homepage: www.sciencedirect.com/journal/journal-of-affective-disorders-reports

Research Paper

The development of a core outcome set for crisis helplines: A three-panel Delphi study

Sonia Curll, Kelly Mazzer *, Debra Rickwood

Discipline of Psychology, Faculty of Health, University of Canberra, 11 Kirinari Street, Bruce, Canberra, Australian Capital Territory 2617, Australia

ARTICLE INFO

Keywords:
Crisis support
Helpline
Suicide
Mental health
Outcome measurement
Delphi
Core outcome set

ABSTRACT

Background: Evidence for the effectiveness of crisis helplines is limited by inappropriate and inconsistent outcome measurement. The aim of this study was to develop a core outcome set that reflects the most relevant and important outcomes to help-seekers accessing a crisis helpline via any delivery mode (e.g., phone, SMS text, online chat).

Method: We used a three-panel Delphi method to compare and integrate the views of three expert groups: people with lived experience of accessing crisis helplines ($n = 32$), researchers with experience assessing crisis helpline outcomes ($n = 25$), and crisis helpline supporters ($n = 58$). Across two online survey rounds (89 % retention rate), participants rated the importance of 33 potential outcomes for help-seekers accessing a crisis helpline. Participants also provided open-text comments and suggestions. Outcomes that reached consensus (≥75 % agreement) by at least two panels were included in the core outcome set.

Results: Ten outcomes met the criteria for inclusion in the core outcome set. In order of importance, these were: *distress, feeling heard, suicide risk, connectedness/support, hopelessness, overwhelm, non-suicidal self-injury risk, service experience, helplessness*, and *next steps*.

Limitations: Participants self-selected and were mainly from English-speaking countries.

Conclusions: We recommend future outcome and evaluation studies minimally measure and report the 10 outcomes identified in this study. Assessing an agreed set of meaningful outcomes will improve comparability and facilitate a deeper understanding of crisis helpline effectiveness. More work is needed to determine *how* best to assess these outcomes in the crisis helpline context.

1. Introduction

Crisis helplines play an important role in mental health service provision and suicide prevention around the world (World Health Organization, 2018). However, little is known about their ability to achieve appropriate outcomes for help-seekers (e.g., Zabelski et al. 2022). To address this knowledge gap, there is a need for more consistent and improved approaches to outcome measurement in crisis helpline research (Mathieu et al., 2021; Mazzer et al., 2020; Trail et al., 2022). The aim of this study was to determine *what* outcomes are most relevant and important to help-seekers accessing a crisis helpline—to inform outcome selection in future research.

For over 60 years, telephone helplines have offered free, immediate, and anonymous support to people experiencing emotional distress or struggling to cope (WHO, 2018). Crisis helplines play a unique role in crisis care internationally (Johnson et al., 2022), giving control to the help-seeker and providing 24/7 support and tailored referrals to other service providers (Pisani et al., 2022; World Health Organization, 2018). In addition to telephone calls, many crisis helplines now offer support via text and online chat, improving accessibility and increasing the demand and diversity of help-seekers (Lake et al., 2022; Lifeline Australia, 2022; Zabelski et al., 2022). As their role expands, a crucial issue is understanding how crisis helplines are meeting users' needs. Studies evaluating crisis helplines tend to report positive outcomes for help-seekers, including reduced risk of suicide, reduced emotional distress, and high service satisfaction (for recent systematic reviews, see Hoffberg et al. 2020, Mazzer et al. 2020). However, the quality of the existing evidence is weak (Hoffberg et al., 2020; Mathieu et al., 2021; Mazzer et al., 2020; Zabelski et al., 2022).

Research on crisis helpline effectiveness is limited by a range of practical and ethical challenges. The anonymous and non-ongoing nature of crisis helpline support makes it difficult to use rigorous follow-up

* Corresponding author.
E-mail address: kelly.mazzer@canberra.edu.au (K. Mazzer).

https://doi.org/10.1016/j.jadr.2024.100763
Received 27 November 2023; Received in revised form 1 March 2024; Accepted 14 March 2024
Available online 15 March 2024

designs (Hvidt et al., 2016; Trail et al., 2022). In addition, the heightened vulnerability of help-seekers raises important ethical issues that limit the use of self-report measures and make the 'gold standard' of randomised controlled trials inappropriate (Britton et al., 2022; Trail et al., 2022). As a result, previous research on crisis helpline outcomes has used a variety of proxy measurement approaches, including crisis supporter assessment, external rater assessment, and analyses of administrative records (Hoffberg et al., 2020). Although such approaches are useful, research assessing the effectiveness of crisis helplines from the perspective of the help-seeker should be prioritised (Gould et al., 2022; Hvidt et al., 2016; Riley et al., 2020; Trail et al., 2022).

The existing literature is further limited by inappropriate and inconsistent outcome measurement. Systematic reviewers have consistently identified a high risk of measurement bias due to the widespread use of unvalidated outcome assessment tools (Hoffberg et al., 2020; Mathieu et al., 2021; Mazzer et al., 2020). Although some studies have used standardised measures, they tend to be mental health screening tools (e.g., Kessler Psychological Distress Scale, General Health Questionnaire) that may not be appropriate for the crisis helpline context (Hvidt et al., 2016; Mazzer et al., 2020; Trail et al., 2022). There has also been wide variation in *what* outcomes have been measured, varying from emotional states (e.g., mood, suicidality, psychological distress, wellbeing) to service satisfaction and referrals. As such, it is difficult to integrate studies and draw firm conclusions about the effectiveness of crisis helpline services (Hoffberg et al., 2020; Mazzer et al., 2020; Zabelski et al., 2022).

To address these issues, researchers have called for the development of measurement tools that are tailored to the crisis helpline setting (Mathieu et al., 2021; Mazzer et al., 2020; Trail et al., 2022). The challenge of defining 'effectiveness' in this context reflects the challenge of the expanding scope expected of crisis helpline services in public health strategies (Zabelski et al., 2022) and in the community (Ma et al., 2022). Indeed, recent research has revealed the diverse functions that crisis helplines can have for help-seekers (Iversen and Westerlund, 2022; Middleton et al., 2017; O'Riordan et al., 2023; Turkington et al., 2020). Although suicide prevention remains the overarching goal, a thorough understanding of the impact of crisis helpline services requires attention to a broader range of suicidal and non-suicidal outcomes (Trail et al., 2022; Zabelski et al., 2022).

A crucial next step in this field of research is to agree *what* outcomes should be measured. To date, outcome selection has been determined by researchers and service providers and hence may be biased toward their perspectives. Incorporating the views of people with lived experience will help ensure that research advances align with what matters most to help-seekers (Hoffberg et al., 2020; Hvidt et al., 2016; Riley et al., 2020).

With increasing reliance on crisis helpline services globally, a more comprehensive understanding of help-seeker outcomes is vital. The development of a core outcome set, based on the views of help-seekers, crisis supporters, and researchers, will help ensure that the most important outcomes are identified and assessed consistently. This will facilitate integration (e.g., meta-analyses) of studies and allow meaningful comparisons across different types of help-seekers, delivery modes (e.g., phone, text, online chat), and helpline service contexts.

2. Current study

The aim of the current study was to identify a core set of outcomes that are relevant and important to help-seekers accessing a crisis helpline via any delivery mode. Developing a core outcome set will inform future research and helpline service design (e.g., program logic models) and evaluation. We used the Delphi method to achieve consensus among three expert groups: (1) people with lived experience of accessing crisis helplines, (2) researchers with experience evaluating crisis helplines, and (3) crisis supporters. This research is reported in accordance with the Core Outcome Set Standards for Reporting (COS-STAR) guidelines (Kirkham et al., 2016).

The Delphi method is a systematic and iterative process for gathering the insights of experts into a group consensus. It assumes that a conclusion based on diverse perspectives will be more valid than a conclusion based on any single expert opinion (Jorm, 2015). A Delphi study involves several stages, including identifying and recruiting relevant experts, constructing survey rounds, administering rounds, analysing rounds, and providing feedback (e.g., Jorm 2015, Jünger et al., 2017). The Delphi method has been increasingly used in healthcare research to identify indicators or outcomes that are important to key stakeholders (Niederberger and Spranger, 2020; Sinha et al., 2011). We used the Delphi method to conduct this research because it provided a rigorous process for integrating knowledge from diverse stakeholders (Beiderbeck et al., 2021).

3. Method

3.1. Expert panel recruitment

We identified three groups with expertise relevant to crisis helpline outcome measurement: (1) people with **lived experience** of using crisis helplines, (2) **researchers** with experience assessing crisis helpline outcomes, and (3) people currently working or volunteering with a crisis helpline as a listening volunteer or **crisis supporter**. We aimed to recruit between 20 and 40 experts for each panel, based on recent recommendations for conducting a Delphi study (Beiderbeck et al., 2021).

Experts were purposively sampled between May and October 2022. Lived experience panelists were recruited via advertisements posted on Lifeline Australia's social media and emailed to all group members of the Lived Experience Advisory Groups for (a) Lifeline Australia, (b) Samaritans U.K., and (c) Roses in the Ocean (Australia's lived experience of suicide organisation). Anyone with lived experience of using a crisis helpline via any delivery mode (phone, text, chat) was eligible to participate. A small payment (equivalent AU$15 per round) was offered to lived experience panelists to acknowledge their unique expertise and in line with research participation guidelines (National Health Medical Research Council, 2019). Researcher panelists were identified via recommendations from project investigators and a literature search of relevant studies published since 2015 in peer-reviewed journals (first and corresponding authors). They were recruited via an email invitation from the research team. Crisis supporter panelists were recruited by (a) Lifeline Australia via an email invitation, and (b) Samaritans U.K. via internal online advertisements. All individuals who currently worked or volunteered as a listening volunteer or crisis supporter for a crisis helpline (including phone, text, chat) were eligible to participate.

Expressions of interest were provided via Qualtrics or email to the research team. All experts that expressed interest were offered to become panel members and were sent a plain language statement that outlined the research aims, what they would be asked to do, and the voluntary and anonymous nature of participation. Experts became panel members on completion of the first survey round.

Ethical approval for the study was provided by the University of Canberra's Human Research Ethics Committee (Project ID: 10377).

3.2. Survey development

We developed an initial outcome pool based on a systematic review of the literature previously conducted by the research team (Mazzer et al., 2020), a targeted search of subsequent literature, and Lifeline Australia's program logic model. Outcomes were retained if they were: (a) help-seeker focused (e.g., not crisis supporter focused), (b) relevant to any help-seeker (e.g., not specific to a demographic group), and (c) relevant to any helpline service delivery mode (including phone call, SMS text, and online chat). Outcome labels and definitions were developed by the research team and then refined by representatives from Lifeline Australia to ensure the language was accessible and

appropriate for the crisis helpline context.

3.3. Survey administration

The Delphi process comprised two online survey rounds. Surveys were administered using Welphi (www.welphi.com). Survey links were distributed via email. Participants provided informed consent at the start of each survey and responses were kept anonymous. Each survey was open for two weeks.

In Round 1, participants were asked to rate the importance of 29 possible outcomes for help-seekers accessing a crisis helpline service via any service delivery mode on a 5-point scale (1=*not important*, 2=*less important*, 3=*important*, 4=*very important*, 5=*essential*). A free-text comment box was available for each outcome. After rating the outcomes, participants could suggest additional outcomes and provide general comments or recommendations. The following demographic data were also collected: gender, age, country of residence, years of experience in research relevant to crisis helplines (researcher panel), years of experience working or volunteering as a crisis supporter (crisis supporter panel), frequency and recency of accessing a crisis helpline (lived experience panel), and service delivery modes provided (crisis supporter panel) or accessed (lived experience panel).

In Round 2, participants were presented with a summary of the Round 1 results (including whether each outcome had been endorsed by any or all of the 3 panels, and the key themes derived from the free-text comments across all panels). They were then asked to re-rate the outcomes that had not yet reached consensus by their panel. Results of the previous round were presented at the time of re-rating so that participants could consider their previous response, as well as the distribution of responses and comments by their panel for each individual outcome. The first point on the response scale was amended to *not important/not realistic* based on participant feedback from Round 1. Participants were also asked to rate four new outcomes that had been suggested by three or more experts in Round 1. Participants were encouraged to provide open-text comments in relation to specific outcomes and general approaches to outcome measurement in the crisis helpline context. The full survey instructions for each round are included in the Supplemental Materials.

3.4. Statistical analysis

Analyses were conducted using Microsoft Excel and SPSS version 25. Results for each panel were first analysed separately and secondly were pooled to give each panel equal weight to account for differences in panel sizes. Aggregate percentages, means, and standard deviations were used to describe the demographic characteristics and responses from each panel after each round. Open-text responses from Round 1 were reviewed and used to refine outcome labels and definitions and introduce new outcomes in Round 2. A summary of the open-text responses was incorporated into the feedback provided to participants after each round. Consensus was defined a priori as $\geq$ 75 % of the panel rating the outcome as very important or essential, based on previous Delphi studies (Jünger et al., 2017). Outcomes that reached consensus by at least two panels were included in the final core outcome set. Outcomes were classified as 'conflicts' based on an analysis of endorsement levels across the three expert panels. Specifically, an outcome was considered in 'conflict' when it achieved consensus within at least one panel and showed an opposing trend with a difference in endorsement of 25 % or more in another panel. This approach allowed us to identify areas of divergence in opinion among the expert groups.

4. Results

4.1. Panel members

A total of 115 experts participated in this study: 32 people with lived experience, 25 researchers, and 58 crisis supporters. There was a strong overall response rate of 89 % in the second round (91 % lived experience, 84 % researchers, 93 % crisis supporters).

Table 1 shows the demographic and expert characteristics of each panel. Most participants were female, with the highest proportion on the lived experience panel (75.9 %). There was a large age range among participants, from 23 to 86 years. The mean age of the lived experience panel was 40 years (ranging from 23 to 71 years), researchers had a mean age of 47 (ranging from 26 to 71), and crisis supporters' mean age was 57 years (ranging from 27 to 86). The majority of participants were from Australia, followed by the United Kingdom, reflecting the targeted recruitment approach. The researcher panel had the highest international representation (6 countries). The high mean years of experience of 13.7 years (researchers) and 5.8 years (crisis supporters) suggests highly experienced samples. Most participants with lived experience had accessed a crisis support service more than six times (61.5 %), and more than three-quarters had accessed a service within the past year (76.0 %).

4.2. Endorsed outcomes

Table 2 shows a summary of the results by panel and round. Six outcomes were endorsed by all three panels: two in Round 1 (*distress, suicide risk*) and four in Round 2 (*feeling heard, hopelessness, overwhelm, non-suicidal self-injury risk*). Four outcomes were endorsed by two panels: three in Round 1 (*helplessness, connectedness/support, service experience*) and one in Round 2 (*next steps*).

The researcher panel endorsed 12 outcomes: *distress, feeling heard, hopelessness, overwhelm, helplessness, hope, connectedness/support, suicide risk, service satisfaction, non-suicidal self-injury risk, next steps*, and *service experience*. They were the only panel to endorse *hope* and *service*

Table 1
Characteristics of the three expert panels.

Characteristic	Lived Experience Panel		Researcher Panel		Crisis Supporter Panel	
	n	%	*n*	%	*n*	%
Gender						
Male	7	24.1	8	36.4	19	33.3
Female	22	75.9	15	63.6	38	66.7
No response	*3*		*2*		*1*	
Country of residence						
Australia	15	55.6	15	62.5	38	76.0
United Kingdom	11	40.7	1	4.2	12	24.0
Other[a]	1	3.7	8	33.3	–	–
No response	*5*		*1*		*8*	
Delivery mode[b]						
Telephone	17	68.0	n/a	n/a	47	100.0
SMS text	9	36.0			12	25.5
Online chat	12	48.0			9	19.1
Email	8	32.0			12	25.5
No response	*7*				*11*	
Access frequency						
Once	4	15.4	n/a	n/a	n/a	n/a
2–5 times	6	23.1				
6–10 times	5	19.2				
11–20 times	2	7.7				
More than 20 times	9	34.6				
No response	*5*					
Access recency						
Within the last 3 months	10	40.0	n/a	n/a	n/a	n/a
3–6 months ago	3	12.0				
6–12 months ago	6	24.0				
1–2 years ago	3	12.0				
More than 2 years ago	3	12.0				
No response	*7*					
Mean age in years (SD)	39.8 (14.0)		46.7 (10.7)		57.0 (14.7)	
Mean years of experience (SD)	n/a		13.7 (8.6)		5.8 (5.6)	

[a] Includes U.S., Taiwan, N.Z., and Denmark.
[b] For the lived experience panel, this refers to the delivery modes accessed. For the crisis supporter panel, this refers to the delivery modes provided. Participants could select multiple response options.

Table 2
Percentage of experts within each panel who endorsed[a] each outcome by round.

Outcome	Lived Experience		Researcher		Crisis Supporter		Result
	Round 1 $n = 32$	Round 2 $n = 29$	Round 1 $n = 25$	Round 2 $n = 21$	Round 1 $n = 58$	Round 2 $n = 54$	
Introduced Round 1:							
Distress	**97**[b]	Consensus[c]	**100**	Consensus	**89**	Consensus	Include[d]
Depression	41	38	8	0	47	37	Exclude
Anxiety	63	65	12	14	62	63	Exclude
Stress	53	45	52	57	59	52	Exclude
Anger	41	41	36	19	43	31	Exclude
Confusion	50	41	48	43	57	54	Exclude
Helplessness	**78**	Consensus	**76**	Consensus	67	72	Include
Overwhelm	**91**	Consensus	72	**76**	69	**80**	Include
Hopelessness	**87**	Consensus	**76**	Consensus	73	**80**	Include
Relief	59	69	44	57	47	50	Exclude
Hope	62	72	72	**76**	53	52	Exclude
Calm	62	65	48	52	56	61	Exclude
Mental Wellbeing	44	41	28	14	24	20	Exclude
Problem Perception	63	66	60	72	42	39	Exclude
Connectedness / Support	72	72	**88**	Consensus	**84**	Consensus	Include
Isolation	**75**	Consensus	60	62	64	70	Exclude
Loneliness	56	65	44	48	41	41	Exclude
Coping Confidence	56	69	56	67	53	54	Exclude
Empowerment	59	69	44	33	55	59	Exclude
Referrals	63	69	60	67	22	19	Exclude
Next Steps	69	**76**	68	**81**	33	32	Include/Conflict[e]
Suicide Risk	**88**	Consensus	**84**	Consensus	**91**	Consensus	Include
Non-Suicidal Self Injury Risk	**78**	Consensus	68	**76**	69	**76**	Include
Other Safety Risk	66	62	60	67	72	74	Exclude
Service Experience	**75**	Consensus	**88**	Consensus	59	48	Include/Conflict
Service Satisfaction	66	66	68	**76**	48	44	Exclude/Conflict
Service Recommendation	41	31	44	43	34	33	Exclude
Willingness to contact again	72	**79**	48	52	48	50	Exclude/Conflict
Introduced Round 2:							
Feeling Heard	–	**100**	–	**76**	–	**98**	Include
Self-Worth	–	48	–	38	–	65	Exclude
Agitation	–	69	–	43	–	67	Exclude
Daily Functioning	–	41	–	29	–	37	Exclude

[a] Endorsed if rated as *very important* or *essential*.
[b] Results shown in bold reflect consensus within the panel ($\geq$ 75 % endorsement).
[c] Outcomes that reached consensus in Round 1 were not re-rated in Round 2.
[d] Included in the core outcome set if 2 or more panels reached consensus.
[e] Outcome was considered a 'conflict' if it reached consensus by at least one panel and there was *a* $\geq$ 25 % difference in% endorsement within another panel.

satisfaction.

The lived experience panel endorsed 11 outcomes: *distress, feeling heard, hopelessness, overwhelm, helplessness, isolation, suicide risk, non-suicidal self-injury risk, next steps, service experience*, and *willingness to contact again*. They were the only panel to endorse *isolation* and *willingness to contact again*.

The crisis supporter panel endorsed seven outcomes: *distress, feeling heard, hopelessness, overwhelm, connectedness/support, suicide risk*, and *non-suicidal self-injury risk*. The lower levels of endorsement among crisis supporters may reflect the larger panel size (Beiderbeck et al., 2021).

4.3. Conflicts

Four outcomes met the definition of a 'conflict' (reached consensus within at least one panel and had a difference of $\geq$ 25 % in endorsement level with another panel): *next steps, service experience, service*

Table 3
Core outcome set: definitions and descriptive statistics.

Outcome	Definition	Lived Experience *M (SD)*	Researcher *M (SD)*	Crisis Supporter *M (SD)*	Overall *GM*	Rank[a]
Distress	Help-seeker's emotional distress has decreased	4.78 (0.49)	4.68 (0.48)	4.37 (0.72)	4.61	1
Feeling Heard	Help-seeker feels heard and validated; that they were listened to without judgement	4.76 (0.44)	4.14 (0.91)	4.83 (0.42)	4.58	2
Suicide Risk	Help-seeker has reduced risk of suicidality (thoughts and/or behaviour)	4.50 (0.76)	4.24 (0.93)	4.52 (0.66)	4.42	3
Connectedness/ Support	Help-seeker experiences a sense of connection and feels supported	4.24 (1.02)	4.36 (0.81)	4.36 (0.79)	4.32	4
Hopelessness	Help-seeker feels less hopeless	4.47 (0.72)	4.00 (1.00)	4.02 (0.71)	4.16	5
Overwhelm	Help-seeker feels less overwhelmed	4.50 (0.67)	3.86 (0.96)	4.06 (0.68)	4.14	6
NSSI Risk	Help-seeker has reduced risk of intentionally hurting themselves	4.00 (1.16)	4.00 (1.05)	4.13 (0.89)	4.04	8
Service Experience	Help-seeker had a positive experience with the service	4.16 (0.99)	4.40 (0.82)	3.56 (1.09)	4.04	8
Helplessness	Help-seeker feels less powerless to improve their situation	4.16 (0.95)	3.96 (0.89)	3.89 (0.77)	4.00	9
Next Steps	Help-seeker has identified positive next step(s) that they can take beyond the contact	4.28 (0.84)	4.19 (0.98)	3.28 (0.98)	3.92	10

[a] Based on the grand mean (mean of means).

satisfaction, and *willingness to contact again* (see Table 2).

4.4. Core outcome set

Ten outcomes reached the criteria to be included in the final core outcome set (i.e., endorsed by at least two panels). In order of importance (based on grand mean rating score), these were: *distress, feeling heard, suicide risk, connectedness/support, hopelessness, overwhelm, non-suicidal self-injury risk, service experience, helplessness,* and *next steps*. Table 3 presents the core outcome set, including the outcome definitions, mean ratings and standard deviations for each panel, and grand mean rating scores.

5. Discussion

This study aimed to identify what outcomes are most relevant and important to help-seekers accessing a crisis helpline. We used a Delphi method drawing on the expertise of people with lived experience of using crisis helplines, researchers with experience evaluating crisis helplines, and crisis helpline service providers. The process yielded 10 key outcomes: *distress, feeling heard, suicide risk, connectedness/support, hopelessness, overwhelm, non-suicidal self-injury risk, service experience, helplessness,* and *next steps*. This core outcome set provides a strong basis for future outcome selection in crisis helpline effectiveness research.

The finding that *distress* and *suicide risk* were among the top three most important outcomes for help-seekers is unsurprising and supports previous work that has focused on these outcomes (e.g., Britton et al. 2022). These outcomes reflect the theoretical underpinnings and overarching goals of most crisis helplines: to prevent suicide by providing immediate and accessible support to individuals experiencing emotional distress (Joiner et al., 2007; Woodward and Wyllie, 2016). Despite the increasing scope expected of 24/7 crisis helplines, these results confirm their key functions to be about reducing distress and suicide risk. It will be important for future work to develop sensitive and appropriate ways to assess suicide-related outcomes in the crisis helpline context (Britton et al., 2022; Hvidt et al., 2016; Sindahl et al., 2019; Trail et al., 2022).

A novel finding in this study was that *feeling heard* was identified as the second most important outcome for help-seekers accessing a crisis helpline. Listening and validating the help-seeker without judgement is recognised as crisis supporter best practice behaviour (Woodward and Wyllie, 2016), and has been used as an indicator of crisis supporter effectiveness (e.g., Gould et al. 2013, Lake et al. 2022, Mishara et al. 2007). However, we are not aware of any studies that have measured the subjective sense of *feeling heard* from a help-seeker perspective. The almost unanimous endorsement among participants with lived experience (100 %) and crisis supporters (98 %) in this study suggests this outcome is key to understanding crisis helpline effectiveness and should be measured in future research.

Beyond general emotional distress, experts agreed on the importance of three specific emotional outcomes for help-seekers: *hopelessness, overwhelm*, and *helplessness*. This finding is consistent with research showing that these feelings are the most frequently expressed by crisis helpline users (Kalafat et al., 2007). These emotional states have also been linked with suicidal ideation (Joiner et al., 2007) and the occurrence of a suicidal crisis (Mishara and Chagnon, 2016; Woodward and Wyllie, 2016), providing further support for their relevance in the crisis helpline context. Recent research conducted in the U.S. has shown that help-seekers tend to feel more hopeful and less overwhelmed after accessing crisis support via SMS text (Gould et al., 2022) and online chat (Gould et al., 2021). Consistent measurement of these specific emotional outcomes may provide a more nuanced understanding of crisis helpline effectiveness and their role in suicide prevention.

It is notable that the outcome of *next steps* was more strongly endorsed than *referrals* across all expert panels. While *referrals* relates specifically to external support pathways, *next steps* can also include adaptive coping strategies (e.g., phoning a friend) and activities associated with daily functioning (e.g., getting dressed). This broader focus may reflect the expanding scope of crisis helplines, including providing out-of-hours support to people with mental disorders, and the need to take a more immediate behavioural activation approach. It may also reflect difficulties in identifying appropriate referrals to meet the diverse needs of help-seekers. We suggest that *next steps* may provide a more meaningful indicator of service effectiveness in linking help-seekers with external support that goes beyond simply providing information (Turley, 2013). However, the crisis supporter panel's low endorsement of next steps (33 %), while higher than for referrals (22 %), reveals a distinct gap in perceived outcome priorities. Promoting a stronger focus among crisis supporters on supporting help-seekers to identify positive short-term next steps may enhance the effectiveness of crisis helplines.

The low endorsement of longer-term outcomes (e.g., *depression, anxiety, loneliness*) supports previous arguments that crisis helpline evaluations should focus on outcomes that are realistic to achieve in a relatively brief and anonymous intervention (Trail et al., 2022; Woodward and Wyllie, 2016). In addition to longer-term outcomes, we also found relatively low endorsement of positively framed outcomes, such as *mental well-being, calm*, and *hope*. Feedback suggested the alleviation of distressing feelings is paramount for help-seekers accessing a crisis helpline, while the achievement of a positive emotional state is considered less realistic.

Our findings point to further potential differences in outcome priorities for help-seekers and service providers. Given the connection between the help-seeker and crisis supporter is fundamental to most service models (Joiner et al., 2007; Woodward and Wyllie, 2016), it is interesting that *connectedness/ support* was not endorsed by the lived experience panel. We speculate that from a user perspective, this connection may be viewed as part of the process, rather than as an outcome in itself of accessing a crisis helpline. A better understanding of the interrelationships between the key outcomes identified in this study may shed light on the pathways of effectiveness in the crisis helpline context. It is also notable that the lived experience panel was the only group to reach consensus on *willingness to contact again* and *isolation*. These results echo recent findings that many help-seekers view crisis helplines as an essential part of their ongoing self-care, which may be at odds with the one-off service model (Iversen and Westerlund, 2022; Middleton et al., 2017). More research is needed to clarify the distinct perspectives identified in this study.

There are clear practical implications of our findings for crisis helpline services. Results provide new insights into the outcomes that are most relevant and important to help-seekers accessing a crisis helpline. This information can be used to inform program logic models and crisis supporter training. Furthermore, the study was able to highlight gaps in the perceptions of crisis supporters and people with lived experience, suggesting priority areas to target to ensure service providers are meeting the needs and expectations of users.

The implications for future research are also considerable. Our results underscore the need to assess a broader range of suicidal and non-suicidal outcomes in crisis helpline effectiveness research (Zabelski et al., 2022). We recommend future studies minimally measure and report the 10 consensus-based outcomes identified in this study. This will reduce the level of uncertainty around outcome selection and allow better integration and comparability of findings across studies. With suicide rates and mental health problems escalating around the world, stronger evidence for the role and impact of crisis helplines is urgently needed.

Important questions remain regarding crisis helpline outcome measurement that were beyond the scope of this study. A crucial next step is determining *how* to measure the outcomes identified in this study. This includes developing and validating outcome measurement tools specific to the crisis helpline context. Innovative data collection approaches that minimise the risks to help-seekers and reduce the burden on crisis supporters should be explored in future research. For example, interactive

voice response systems could be used to present outcome measures to help-seekers and allow them to respond using numerical telephone keys (Mathieu et al., 2021; Trail et al., 2022). Developing unobtrusive ways to determine outcomes, such as through computational linguistics and natural language processing, may also be valuable. Finally, there is a need to better understand the longer-term impacts of crisis intervention (Zabelski et al., 2022), including identifying what distal outcomes matter most to help-seekers. While we encourage future research in this area, our findings suggest experts believe it may be unrealistic to expect longer-term impacts of a crisis helpline on mental health outcomes (e.g., depression, anxiety), and that post-contact outcomes relating to positive next steps and willingness to contact again may be more important to help-seekers.

5.1. Strengths and limitations

A strength of this study was the use of the Delphi method. Our three-panel approach allowed us to incorporate three types of critical expertise, while maintaining participant anonymity and avoiding power imbalances and personal sensitivities that had the potential to undermine the consensus process (Sinha et al., 2011). By surveying and analysing each panel separately, we were able to explore similarities and differences between them. The online survey method allowed experts from geographically diverse locations to participate. The high response rate across the survey rounds demonstrates the engagement of the expert participants and supports the validity of the results (Sinha et al., 2011).

A main limitation of this study was the generalisability of the results. Participants self-selected and our purposive sampling approach focused on people with lived experience and crisis supporters associated with two large national helplines (Lifeline Australia and Samaritans U.K.). Hence, there could be bias in the samples of help-seekers, crisis supporters, and researchers that participated, including a lack of representation from low- and middle- income countries and non-English speaking participants. Further, the online nature of this study might have restricted the participation of other vulnerable groups with complex needs (e.g., people experiencing homelessness, people with a disability). It is also important to consider that this study focused on identifying the key outcomes relevant to crisis helplines; future investigations are needed to determine if these key outcomes also apply to other types of crisis support services (e.g., crisis cafés; Johnson et al., 2022). We encourage researchers to continue to explore additional outcomes that may be important to specific helplines, delivery modes, or types of help-seekers (e.g., youth; Tibbs et al., 2022).

6. Conclusion

Although there is increasing reliance on crisis helplines in suicide prevention and public health provision, important knowledge gaps remain regarding user outcomes. This study advances our knowledge by identifying a core set of outcomes that reflect the perspectives of three critical stakeholder groups: researchers, crisis supporters, and the service users themselves. We recommend researchers and service providers use the results of this study to guide outcome selection in future crisis helpline evaluations. Assessing an agreed set of person-centred outcomes will facilitate a deeper understanding of the role and impact of crisis helplines globally. Future work is needed to determine *how* best to measure these outcomes within the crisis helpline context.

Funding

This work was conducted by the University of Canberra as part of a National Health and Medical Research Council (NHMRC) Partnership Grant with Lifeline Australia (GNT1153481).

Role of the funding source

The NHMRC had no involvement in the conduct of this research or the preparation of this article.

Publication ethics

This work was approved by the University of Canberra's Human Research Ethics Committee (Project ID: 10377). The study was conducted in accordance with ethical standards of the University of Canberra's Human Research Ethics Committee. Informed consent was obtained from all participants included in the study.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Sonia Curll: Methodology, Project administration, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Kelly Mazzer:** Writing – review & editing, Supervision, Investigation, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Debra Rickwood:** Writing – review & editing, Supervision, Software, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

None

Acknowledgments

We would like to thank all panel members for their time, expertise, and ongoing commitment to this project. We also thank Lifeline Australia and Lifeline Centres, Samaritans U.K., and Roses in the Ocean for their contributions to recruitment and support for this project.

References

Beiderbeck, D., Frevel, N., Heiko, A., Schmidt, S.L., Schweitzer, V.M., 2021. Preparing, conducting, and analyzing Delphi surveys: cross-disciplinary practices, new directions, and advancements. MethodsX 8, 101401. https://doi.org/10.1016/j.techfore.2021.120577.

Britton, P.C., Karras, E., Stecker, T., Klein, J., Crasta, D., Brenner, L.A., Pigeon, W.R., 2022. Veterans crisis line call outcomes: distress, suicidal ideation, and suicidal urgency. Am. J. Prev. Med. 62, 745–751. https://doi.org/10.1016/j.amepre.2021.11.013.

Gould, M.S., Chowdhury, S., Lake, A.M., Galfalvy, H., Kleinman, M., Kuchuk, M., McKeon, R., 2021. National Suicide Prevention Lifeline crisis chat interventions: evaluation of chatters' perceptions of effectiveness. Suicide Life Threat. Behav. 51, 1126–1137. https://doi.org/10.1111/sltb.12795.

Gould, M.S., Cross, W., Pisani, A.R., Munfakh, J.L., Kleinman, M., 2013. Impact of applied suicide intervention skills training on the national suicide prevention lifeline. Suicide Life Threat. Behav. 43, 676–691. https://doi.org/10.1111/sltb.12049.

Gould, M.S., Pisani, A., Gallo, C., Ertefaie, A., Harrington, D., Kelberman, C., Green, S., 2022. Crisis text-line interventions: evaluation of texters' perceptions of effectiveness. Suicide Life Threat. Behav. 52, 583–595. https://doi.org/10.1111/sltb.12873.

Hoffberg, A.S., Stearns-Yoder, K.A., Brenner, L.A., 2020. The effectiveness of crisis line services: a systematic review. Front. Public Health 7, 399. https://doi.org/10.3389/fpubh.2019.00399.

Hvidt, E., Ploug, T., Holm, S., 2016. The impact of telephone crisis services on suicidal users: a systematic review of the past 45 years. Ment. Health Rev. 21, 141–160. https://doi.org/10.1108/MHRJ-07-2015-0019.

Iversen, C., Westerlund, M., 2022. Users' perspectives on crisis helplines in relation to professional mental health services. Crisis. https://doi.org/10.1027/0227-5910/a000876. Advance online publication.

Johnson, S., Dalton-Locke, C., Baker, J., Hanlon, C., Salisbury, T.T., Fossey, M., Lloyd-Evans, B., 2022. Acute psychiatric care: approaches to increasing the range of

services and improving access and quality of care. World Psychiatry 21, 220–236. https://doi.org/10.1002/wps.20962.
Joiner, T., Kalafat, J., Draper, J., Stokes, H., Knudson, M., Berman, A.L., McKeon, R., 2007. Establishing standards for the assessment of suicide risk among callers to the National Suicide Prevention Lifeline. Suicide Life Threat. Behav. 37, 353–365. https://doi.org/10.1521/suli.2007.37.3.353.
Jorm, A.F., 2015. Using the Delphi expert consensus method in mental health research. Aust. N. Z. J. Psychiatry 49, 887–897. https://doi.org/10.1177/0004867415600891.
Jünger, S., Payne, S.A., Brine, J., Radbruch, L., Brearley, S.G., 2017. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat. Med. 31, 684–706. https://doi.org/10.1177/0269216317690685.
Kalafat, J., Gould, M.S., Munfakh, J.L.H., Kleinman, M., 2007. An evaluation of crisis hotline outcomes. Part 1: nonsuicidal crisis callers. Suicide Life Threat. Behav. 37, 322–337. https://doi.org/10.1521/suli.2007.37.3.322.
Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Williamson, P.R., 2016. Core outcome set–STAndards for reporting: the COS-STAR statement. PLoS Med. 13, e1002148 https://doi.org/10.1371/journal.pmed.1002148.
Lake, A.M., Niederkrotenthaler, T., Aspden, R., Kleinman, M., Hoyte-Badu, A.M., Galfalvy, H., Gould, M.S., 2022. Lifeline Crisis Chat: coding form development and findings on chatters' risk status and counselor behaviors. Suicide Life Threat. Behav. 52, 452–466. https://doi.org/10.1111/sltb.12835.
Lifeline Australia, 2022. Annual Report: 2021-2022. Australia. https://www.lifeline.org.au/media/hikdmw5w/lifeline-annual-report-2022.pdf.
Ma, J.S., Batterham, P.J., Kõlves, K., Woodward, A., Bradford, S., Klein, B., Rickwood, D. J., 2022. Community expectations and anticipated outcomes for crisis support services—Lifeline Australia. Health Soc. Care Community 30, 1775–1788. https://doi.org/10.1111/hsc.13557.
Mathieu, S.L., Uddin, R., Brady, M., Batchelor, S., Ross, V., Spence, S.H., Watling, D., Kõlves, K., 2021. Systematic review: the state of research into youth helplines. J. Am. Acad. Child Adolesc. Psychiatry 60, 1190–1233. https://doi.org/10.1016/j.jaac.2020.12.028.
Mazzer, K., O'Riordan, M., Woodward, A., Rickwood, D., 2020. A systematic review of user expectations and outcomes of crisis support services. Crisis 42, 465–473. https://doi.org/10.1027/0227-5910/a000745.
Middleton, A., Woodward, A., Gunn, J., Bassilios, B., Pirkis, J., 2017. How do frequent users of crisis helplines differ from other users regarding their reasons for calling? Results from a survey with callers to Lifeline, Australia's national crisis helpline service. Health Soc. Care Community 25, 1041–1049. https://doi.org/10.1111/hsc.12404.
Mishara, B.L., Chagnon, F., 2016. Why mental illness is a risk factor for suicide: implications for suicide prevention. O'Connor, R., Pirkis, J. (Eds.). The International Handbook of Suicide Prevention. Wiley-Blackwell, Chichester, pp. 594–608.
Mishara, B.L., Chagnon, F., Daigle, M., Balan, B., Raymond, S., Marcoux, I., Berman, A., 2007. Comparing models of helper behavior to actual practice in telephone crisis intervention: a silent monitoring study of calls to the US 1-800-SUICIDE network. Suicide Life Threat. Behav. 37, 291–307. https://doi.org/10.1111/sltb.1138.
National Health Medical Research Council, 2019. Payment of participants in research: information for researchers, HRECs and other ethics review bodies. National Health and Medical Research Council, Australian Research Council and Universities Australia. Commonwealth of Australia, Canberra. https://www.nhmrc.gov.au/about-us/publications/payment-participants-research-information-researchers-hrecs-and-other-ethics-review-bodies.
Niederberger, M., Spranger, J., 2020. Delphi technique in health sciences: a map. Front. Public Health 8, 457. https://doi.org/10.3389/fpubh.2020.00457.
O'Riordan, M., Rickwood, D., Curll, S., 2023. What is a crisis? Perspectives of crisis support help-seekers. Crisis. https://doi.org/10.1027/0227-5910/a000910. Advance online publication.
Pisani, A.R., Gould, M.S., Gallo, C., Ertefaie, A., Kelberman, C., Harrington, D., Green, S., 2022. Individuals who text crisis text line: key characteristics and opportunities for suicide prevention. Suicide Life Threat. Behav. 52, 567–582. https://doi.org/10.1111/sltb.12872.
Riley, J., Mok, K., Larsen, M., Boydell, K., Christensen, H., Shand, F., 2020. Meeting the needs of those in suicidal crisis with new models and integrated care, in: Black Dog Institute, What can be done to decrease suicidal behaviour in Australia? A call to action. Sydney, Australia. https://www.blackdoginstitute.org.au/wp-content/uploads/2020/09/What-Can-Be-Done-To-Decrease-Suicide_Chapter-1-New-Models-of-Care.pdf.
Sindahl, T.N., Côte, L.P., Dargis, L., Mishara, B.L., Bechmann Jensen, T., 2019. Texting for help: processes and impact of text counseling with children and youth with suicide ideation. Suicide Life Threat. Behav. 49, 1412–1430. https://doi.org/10.1111/sltb.12531.
Sinha, I.P., Smyth, R.L., Williamson, P.R., 2011. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Med. 8, e1000393 https://doi.org/10.1371/journal.pmed.1000393.
Tibbs, M., O'Reilly, A., O'Reilly, M.D., Fitzgerald, A., 2022. Online synchronous chat counselling for young people aged 12–25: a mixed methods systematic review protocol. BMJ Open 12, e061084. https://doi.org/10.1136/bmjopen-2022-061084.
Trail, K., Baptiste, P.J., Hunt, T., Brooks, A., 2022. Conducting research in crisis helpline settings. Crisis 43, 263–353. https://doi.org/10.1027/0227-5910/a000858.
Turkington, R., Mulvenna, M.D., Bond, R.R., O'Neill, S., Potts, C., Armour, C., Ennis, E., Millman, C., 2020. Why do people call crisis helplines? Identifying taxonomies of presenting reasons and discovering associations between these reasons. J. Health Inform. 26, 2597–2613. https://doi.org/10.1177/1460458220913429.
Turley, B., 2013. Crisis support: The legacy and Future of Helplines. Lifeline Foundation Australia. https://www.lifeline.org.au/media/j12j5jmt/crisis-support-the-legacy-and-future-of-helplines-2013.pdf.
Woodward, A., Wyllie, C., 2016. Helplines, tele-web support services, and suicide prevention. O'Connor, R., Pirkis, J. (Eds.). The International Handbook of Suicide Prevention. Wiley-Blackwell, Chichester, pp. 490–504.
World Health Organization, 2018. Preventing suicide: a resource for establishing a crisis line. Geneva, Switzerland. https://apps.who.int/iris/bitstream/handle/10665/311295/WHO-MSD-MER-18.4-eng.pdf.
Zabelski, S., Kaniuka, A.R., Robertson, A.R., Cramer, R.J., 2022. Crisis lines: current status and recommendations for research and policy. Psychiatr. Serv. 74, 505–512. https://doi.org/10.1176/appi.ps.20220294.