

## Short report

# Effect of feedback on the quality of suicide prevention websites: randomised controlled trial

Anthony F. Jorm, Julie-Anne Fischer and Elizabeth Oh

## Summary

There is concern regarding the quality of information about mental health problems on the internet. A trial was carried out to see whether sending feedback to website administrators about the quality of information on their website would lead to an improvement (ACTRN12609000449235). Fifty-two suicide prevention websites were identified by means of an online search. The quality of information about how to help someone

who is suicidal was scored against expert consensus guidelines. Websites were randomised to receive feedback or serve as controls. The information on the websites varied greatly in quality. However, feedback did not lead to an improvement.

## Declaration of interest

None.

The internet is a major source of information for the public about health problems, including mental health problems.<sup>1</sup> However, formal evaluations of information on websites have frequently noted that the quality is low or variable.<sup>2–7</sup> Standards have been proposed for website quality,<sup>8</sup> but these have not affected the widespread availability of low-quality information. Here we report what we believe to be the first randomised controlled trial of an intervention to improve the quality of websites, using suicide prevention websites as the target. Little information is available on use of the internet to assist someone who is suicidal. However, many members of the Australian public believe that website information is potentially helpful to a suicidal person,<sup>9</sup> and in the UK the Samaritans report receiving many requests for help through email.<sup>10</sup>

## Method

Our sample comprised the top 52 English-language websites on suicide prevention identified from a Google Australia search using the terms 'help suicide'. The search was not meant to be exhaustive, but rather to locate sites that might be found by a member of the public looking for assistance. Website content was captured on 6 November 2008, with the use of TeleportPro software version 1.60 ([www.tenmax.com](http://www.tenmax.com)). Websites were included if they were sponsored by any organisation (to exclude those maintained by an individual, such as a blog) and listed the contact details of a website administrator.

Websites were scored for quality of information by two of the authors (J-A.F. and E.O.) using a 26-item checklist, giving one point per suicide prevention action recommended, based on the document *Suicidal Thoughts and Behaviours: First Aid Guidelines*, which was developed through a Delphi expert consensus study.<sup>11,12</sup> Administrators of the websites allocated to receive the intervention were emailed a letter from A.F.J. giving detailed scoring on the checklist and the mean score of all 52 websites for comparison. The letter was headed 'University of Melbourne Project to Evaluate and Improve Suicide Prevention Websites' and had a large university logo at the top to add authority. The administrators were also sent a copy of the guidelines in the form of an attractively designed portable file that could be easily posted, an article on how the guidelines were produced,<sup>12</sup> and an information sheet about the study. This sheet explained how the websites were selected, the nature of the intervention, the random assignment to intervention or control groups, and when the post-test assessment would occur. The control websites were not communicated with at all and

received no feedback. Websites were scored at baseline (in months 1 and 2 of the study) and 6 months later (in months 6 and 7). Intervention letters were sent out 2 months after baseline scoring. This allowed a period of 4 months for website administrators to make any changes in response to the feedback.

The outcome was the quality of information about how to assist someone who is suicidal. Quality was measured against a 26-item checklist. Each item was scored as 'yes' or 'no', depending on whether or not a particular action was recommended on the website. Two of the authors (J-A.F. and E.O.) independently selected content from each website and came to a consensus about which text was relevant to score. They also scored the material independently and came to a consensus about the quality score at both assessment points.

There were 52 websites. This sample size gives 80% power to detect a large effect size ( $d=0.8$ ) at  $\alpha=0.05$ .<sup>13</sup> To put this effect size in concrete terms, it would have been achieved if 7 out of the 26 intervention group websites had simply posted the guidelines on their website and hence received the maximum score of 26.

Websites were given an identification number (1–52) and an independent researcher used the Random.org sequence generator ([www.random.org](http://www.random.org)) to randomly order these. The first 26 numbers were allocated to the feedback group and the second 26 to the control group. These random allocations were provided to A.F.J. who linked them to a list of websites, their web addresses and quality scores provided by J-A.F. Feedback letters containing the quality scores and guidelines for improvement were emailed to website administrators by A.F.J. Neither of the researchers involved in scoring the websites (J-A.F. and E.O.) knew the condition to which the websites were randomised. The trial was registered with the Australia New Zealand Clinical Trials Registry (ACTRN12609000449235). Ethical approval was provided by the University of Melbourne ethics committee.

Interrater reliability and stability of quality ratings were assessed with Pearson correlations. The effect of the intervention on quality scores was assessed using analysis of variance, where the factors were intervention (feedback or control) and time (pre-test or post-test). The  $P<0.05$  significance level was used.

## Results

Interrater reliability was high, with correlations of 0.90 at pre-test and 0.92 at post-test. Quality was also stable over time, with a correlation between the pre-test and post-test consensus ratings

of 0.79. The quality of the information provided was highly variable, with pre-test scores ranging from 1 to 19 with a mean of 9.79 (s.d.=4.18). The two most highly rated websites both before and after the intervention were Suicide.org (<http://www.suicide.org>) and Suicideline (<http://suicideline.org.au>).

Online Fig. DS1 shows the CONSORT flow diagram for the trial. There was no significant difference in mean quality scores at pre-test ( $t=0.96$ , d.f.=50,  $P=0.34$ ). The mean scores of the intervention websites changed from 9.23 (s.d.=4.85, 95% CI 7.58–10.88) at pre-test to 11.62 (s.d.=4.81, 95% CI 9.84–13.39) at post-test. For the control websites the respective means were 10.35 (s.d.=3.38, 95% CI 8.70–11.99) and 11.54 (s.d.=4.16, 95% CI 9.77–13.31). Analysis of variance showed a significant effect of time ( $F=21.31$ , d.f.=1,50,  $P<0.001$ ), but no difference between intervention and control websites ( $F=0.21$ , d.f.=1,50,  $P=0.65$ ) or intervention time interaction ( $F=2.37$ , d.f.=1,50,  $P=0.13$ ). A *post hoc* examination of whether a website improved by 5 points or more (corresponding to roughly a 25% improvement or more than 1 standard deviation unit) showed that five intervention websites improved to this degree compared with one control website, but this difference was not significant (two-tailed  $P=0.19$ , Fisher's exact test). Although no reply was requested, administrators of nine websites wrote personal replies acknowledging the feedback or commenting on it. *Post hoc* analysis showed that the intervention had no greater impact on these websites than on the other intervention or control websites ( $P=0.30$  for interaction effect).

## Discussion

Our findings are consistent with most previous evaluations of website information in finding many sites of low quality. However, we found no evidence that feedback to website administrators had any effect. One possible reason is that the feedback did not reach an appropriate person: emails to websites might not be read at all or read by an information technology employee who takes no interest in the content. Although a *post hoc* analysis of the data did not support this interpretation, the number of websites from which we received correspondence was small. This trial can be seen as an effectiveness rather than an efficacy trial, because it evaluated feedback under realistic conditions, but it does not tell us whether or not feedback would work if it was actually received by the relevant person. Another possibility is lack of statistical power: we set the sample size to detect a large effect size on the basis that the intervention would lead to a large improvement in quality score. Furthermore, a feedback intervention resulting in small effects is of little practical relevance. However, a *post hoc* analysis of power, using the observed pre-test post-test correlation in analysis of covariance, showed that the actual power was greater than planned, with 81% power to detect a medium effect size ( $\alpha=0.05$ , two-tailed). Power to detect a small effect size was low (21%).

Although the intervention had no effect, there was a significant overall improvement from pre-test to post-test. Real improvement might have occurred independently of the intervention, but we believe that the change was mainly due to greater experience of the raters, both of whom felt that they had become more skilled at finding relevant information on the websites, because they had already scored and discussed them. However, given that the raters were masked to allocation, this improvement did not affect the validity of the comparison of intervention and control websites. It may have meant that the pre-test score used for feedback was an underestimate of a website's quality, but even

after the intervention all websites had room for improvement, most substantially so, which was the main message the feedback conveyed.

Given the lack of success of this simple intervention, what else could be done to improve website quality? One possibility is to engage in a dialogue with the organisations behind the websites, including exploration of barriers to making changes. This might include personal contact by telephone. There may be useful guidance on how to produce behaviour change from the general literature on computerised health interventions.<sup>14</sup> Some of the website administrators requested further information, such as examples of better-quality websites; however, the protocol set up and registered for this trial was to investigate the impact of simple feedback alone and we did not deviate from this. We hope that this first attempt at a randomised trial in the area of website quality may influence other researchers to look at more sophisticated interventions.

**Anthony F. Jorm, PhD, DSC, Julie-Anne Fischer, BA, Elizabeth Oh, BSc,** Orygen Youth Health Research Centre, Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia

**Correspondence:** Anthony F. Jorm, Orygen Youth Health Research Centre, Locked Bag 10, Parkville, Victoria 3052, Australia. Email: ajorm@unimelb.edu.au

First received 8 Sep 2009, final revision 11 Feb 2010, accepted 30 Mar 2010

## Acknowledgements

Thanks to Amy Morgan for assistance. Support was provided by the National Health and Medical Research Council and the Colonial Foundation.

## References

- 1 Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. *Cyberpsychol Behav* 2004; **7**: 497–510.
- 2 Ipser JC, Dewing S, Stein DJ. A systematic review of the quality of information on the treatment of anxiety disorders on the internet. *Curr Psychiatry Rep* 2007; **9**: 303–9.
- 3 Griffiths KM, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. *BMJ* 2000; **321**: 1511–5.
- 4 Khazaal Y, Fernandez S, Cochand S, Reboh I, Zullino D. Quality of web-based information on social phobia: a cross-sectional study. *Depress Anxiety* 2008; **25**: 461–5.
- 5 Murphy R, Frost S, Webster P, Schmidt U. An evaluation of web-based information. *Int J Eat Disord* 2004; **35**: 145–54.
- 6 Khazaal Y, Chatton A, Cochand S, Jermann F, Osiek C, Bondolfi G, et al. Quality of web-based information on pathological gambling. *J Gambl Stud* 2008; **24**: 357–66.
- 7 Kisely S, Ong G, Takyar A. A survey of the quality of web based information on the treatment of schizophrenia and attention deficit hyperactivity disorder. *Aust N Z J Psychiatry* 2003; **37**: 85–91.
- 8 Commission of the European Communities. eEurope 2002: quality criteria for health related websites. *J Med Internet Res* 2002; **4**: e15.
- 9 Leach LS, Christensen H, Griffiths KM, Jorm AF, Mackinnon AJ. Websites as a mode of delivering mental health information: perceptions from the Australian public. *Soc Psychiatry Psychiatr Epidemiol* 2007; **42**: 167–72.
- 10 Dobson R. Internet sites may encourage suicide. *BMJ* 1999; **319**: 337.
- 11 Mental Health First Aid. *Suicidal Thoughts and Behaviours: First Aid Guidelines* ([http://www.mhfa.com.au/documents/guidelines/8191\\_MHFA\\_suicide\\_guidelines.pdf](http://www.mhfa.com.au/documents/guidelines/8191_MHFA_suicide_guidelines.pdf)).
- 12 Kelly CM, Jorm AF, Kitchener BA, Langlands RL. Development of mental health first aid guidelines for suicidal ideation and behaviour: a Delphi study. *BMC Psychiatry* 2008; **8**: 17–27.
- 13 Cohen J. A power primer. *Psychol Bull* 1992; **112**: 155–9.
- 14 Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. *Patient Educ Couns* 2009; **74**: 156–73.

