Suicide prevention is everyone's business: Challenges and opportunities for Google

Olivia J. Kirtley\textsuperscript{a,}\textsuperscript{*}, Rory C. O'Connor\textsuperscript{b}

\textsuperscript{a} Center for Contextual Psychiatry, KU Leuven, Department of Neuroscience, Campus Sint-Rafael, Kapucijnenvoer 33, Bus 7001 (Blok H), 3000, Leuven, Belgium

\textsuperscript{b} Suicidal Behaviour Research Laboratory, Institute of Health & Wellbeing, University of Glasgow, Mental Health & Wellbeing, Academic Centre, Gartnavel Royal Hospital, 1055 Great Western Road, Glasgow, G12 0XH, UK

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Suicide prevention
Internet
Public health
Intervention

\textbf{ABSTRACT}

The internet has become a key frontier for large-scale, public health efforts in suicide prevention. Market-leading technology companies, such as Google, are developing interventions to deliver support information to those experiencing suicidal distress, but the precise technology, i.e. algorithms, behind this are proprietary. This raises important ethical questions regarding whether such large-scale public health interventions for suicide prevention should be happening behind closed corporate doors when this makes the evaluation of such interventions extremely difficult. Furthermore, as illustrated by Arendt et al. (2019), initiatives such as Google's Suicide Prevention Result (SPR) appear not to work in circumstances in which they could be of significant potential benefit, such as when individuals are searching for details of celebrity suicides. In the current commentary, we discuss ways in which the SPR can be optimized, based on existing evidence regarding suicide-related internet use. We go on to discuss the ethical issues of large technology companies becoming key players in suicide prevention and critically consider how online public health initiatives of this kind are able to be evaluated.

\textbf{1. Introduction}

Suicide is a global public health issue, with more than 800,000 people dying by suicide globally each year (World Health Organization, 2014). As such, suicide needs addressing not only at the micro (individual) level but also at the macro (population) level via large-scale, “universal” public health interventions. Recent figures show that over 4.33 billion people worldwide are active internet users (We Are Social, & DataReportal and Hootsuite, 2019) and that Asia is ahead of North America and Europe with 2.6 billion internet users compared to 0.3 billion and 0.7 billion users, respectively (Internet World Stats, 2018). Thus, for suicide prevention on a global scale, few means present a better opportunity for intervention on such a vast scale than harnessing the reach of the internet. Aside from the relative ubiquity of access, the internet is an important target for suicide prevention efforts, as previous research has demonstrated associations between exposure to suicide-related content online and having a lifetime history of suicidal ideation and behaviors (Marchant et al., 2017; Mars et al., 2015; O’Connor et al., 2014). How then can one use the internet as a suicide prevention tool? In recent years, a market-leading search engine company, Google, has developed the “Suicide Prevention Result” (SPR), where in 14 countries, contact details for the relevant national suicide crisis line are displayed when users search for particular keywords around the topic of suicide (Zeiger, 2010).

The article by Arendt et al. (2019a), highlights a number of important issues regarding the SPR. First, that the SPR appears to be operated via a somewhat crude keyword search, meaning that crucial opportunities for suicide prevention are missed. Namely, when individuals’ searches include the name of a celebrity, as well as suicide-related keywords, the SPR is not displayed across four German-speaking countries. This finding is particularly stark when considered alongside the evidence that the SPR was displayed in about half of all searches in two of the countries studied when suicide-related terms alone were searched.

Second, Arendt et al. (2019a), made use of agent-based testing techniques, whereby a computer program imitates individuals performing particular actions, in this case Google searches for suicide-related content. Their innovation was, however, a necessity as the algorithm that underpins Google’s SPR is proprietary. This practice raises ethical questions regarding whether or not the technology underlying such a large-scale public health intervention for preventing suicide should be a black box, hidden behind closed corporate doors. Moreover, the technology is implemented differently across the globe (Scherr et al., 2019), and again it is not clear what legal and commercial
constraints drive this activity. Third, given that the SPR algorithm is proprietary and inaccessible to those outside of Google, how can this intervention be effectively evaluated? Indeed, how should such large-scale, internet-based public health initiatives be evaluated? In the current commentary, we will discuss challenges and opportunities for optimizing Google’s SPR, then go on to examine the ethical issues that arise when suicide prevention becomes proprietary business and finally, whether and how such interventions can be evaluated.

1.1. Optimizing the SPR: opportunities and challenges

Arendt et al. (2019a), demonstrated that, in four German-speaking nations, the SPR is not displayed when people use suicide-related search terms alongside the names of celebrities. Given evidence that suggests that some individuals do specifically search for names of celebrities who have died by suicide in order to find suicide methods details (Biddle et al., 2018), this practice amounts to a concerning “Achilles heel” within the SPR. This issue is important, as recent research has reported a 10% excess in deaths by suicide in the general population in the months following a celebrity suicide (Pink et al., 2018). Still, as noted by Stack’s (in press) commentary, there is considerable debate about the extent and nature of the association between celebrity suicides and suicide rates in general. There is also discussion about the extent of the relationship between internet searches for suicide-related content and celebrity suicides. A recent study used Google Trends data to investigate searches for “how to suicide” and “suicide prevention” in the weeks before and after 16 celebrity suicides, but found increased search volumes only for two celebrities: Robin Williams and Aaron Hernandez (Gunn et al., 2019). Irrespective of the volume of searches, the content of the search results may change after, relative to before, a celebrity dies by suicide. Google Trends does not provide information regarding content of search results, however future research should attempt to take a more nuanced and wide-ranging approach to understanding how search terms and results for suicide-related (and celebrity-related) content may vary temporally in relation to celebrity suicide.

Nonetheless, in order to meaningful function as a suicide prevention intervention or tool, it is vital that the SPR incorporates high quality research evidence about the nature of suicide-related internet use to avoid missed opportunities such as revealed in Arendt et al. (2019a). As suicides can be prevented practically until the final moment, it is vital that all possible efforts are made to protect the most vulnerable in society. Indeed contemporary theoretical models of suicide, like the integrated motivational-volitional model of suicidal behaviour (IMV; O’Connor and Kirtley, 2018) emphasize the importance of targeting the volitional phase of the suicidal process to reduce the likelihood that individuals who are acutely suicidal will act on their thoughts. Volitional phase factors are defined as those factors that facilitate or impede the transition from suicidal thoughts to suicide attempts. For the present purposes, SPRS may interrupt suicidal planning by increasing the likelihood that if an individual feels suicidal then they will contact a crisis helpline rather than acting on their thoughts of suicide. Moreover, information from the broader suicide research literature can also potentially help to refine and extend the SPR.

The precise mechanism through which exposure to suicidal behaviour (via any means) influences individuals’ own suicidal behaviour is poorly understood, yet, models from the non-suicidal self-injury (NSSI) literature, such as the Cognitive-Emotional model of NSSI (Hasking et al., 2017), have begun to explore this very issue. Hasking et al. (2017) describe how exposure to NSSI can alter individuals’ outcome expectancies about engaging in the behaviour by, for example, reducing how painful people think NSSI will be or increasing the extent to which people think NSSI will relieve their emotional pain. Others have found that if someone else engages in NSSI, it may legitimize the behaviour or make it more cognitively accessible (O’Connor et al., 2014). While the Cognitive-Emotional model (Hasking et al., 2017) was developed in the context of NSSI, some of these outcome expectancies may also have relevance for suicide. For example, Biddle and colleagues’ (2010) qualitative study found that individuals admitted to hospital following a suicide attempt by hanging had specifically chosen that method due to perceiving it to be “simple”, “rapid” and “painless”, whereas others who used other means of trying to end their life had actively rejected hanging due to having negative outcome expectancies regarding the method. A potential future target for the SPR, therefore, could be to display in response to searches involving outcome expectancies for suicide and specific suicide methods, such as searches for painfulness, easiness or expected lethality of suicide.

While the names of celebrities who have died by suicide could be added to the keyword list that feeds the SPR algorithm, an additional consideration is whether the names of fictional characters who died by suicide or TV shows in which a suicide occurs should also be added to the list. Previous research by Ayers et al. (2017) demonstrated a significant increase in suicide-related internet searches (including for help-seeking purposes) following the release of the first season of the Netflix series, 13 Reasons Why. The show attracted widespread concern from the suicide prevention community for an extremely graphic depiction of the main protagonist’s suicide (Arensman et al., 2017) together with calls of restraint to ensure that we do not shut down conversations with young people about suicide (O’Connor, 2017). The concern was warranted, though, as illustrated by a recent study by Niederkrotenthaler et al. (2019) who found a 21.7% increase in suicide deaths amongst 10-19 year-old females (the main target demographic of 13 Reasons Why), which correlated with the period of greatest interest around the show, as indicated by social media traffic.

Niederkrotenthaler and colleagues urge caution in interpreting the results of their study, as they are correlational and from ecological data, meaning it is not known whether or not those individuals who died by suicide had actually watched the show. Moreover, because the expected rise in suicide by cutting was not observed; suicide by hanging increased, but the frequency of suicide by cutting was too low to permit time-series analysis. Evidence around whether dramatic depictions of suicide and suicide attempts are associated with increases in suicide attempts and deaths is mixed, as discussed more extensively in Stack’s (in press) commentary on Arendt et al. (2019a) study, although some earlier evidence suggests an increase in suicidal behaviour using the same methods as depicted in another popular television show, immediately following its broadcast (Hawton et al., 1999). A key criticism of research on exposure to suicide is that the majority of studies cannot demonstrate that the individuals who attempt or die by suicide have actually been exposed to the content in question. Arendt et al., (2019b) recently addressed this in a longitudinal panel study of young adults, by conducting baseline assessments prior to the release of season 2 of series 13 Reasons Why and then a follow-up one month later. Crucially, Arendt et al. (2019b) also considered additional characteristics of participants beyond their exposure to 13 Reasons Why. Their findings demonstrate that, on one hand, watching the show was harmful for some individuals – students and those who did not watch the full season 2. On the other hand, the show appeared to have beneficial effects for others – non-students and those watching the complete season 2. The suicide of fictional characters may, therefore, not speak similarly or equally to all demographics (Mueller, 2019), which may mean that an SPR displaying in response to searches for fictional characters could appeal only to specific groups of internet users.

Media guidelines regarding the reporting of suicide have been developed by a number of organizations including the International Association for Suicide Prevention in conjunction with the World Health Organization (IASP, 2017), Samaritans (2013) and Everymind (2014), and the first guidelines for filmmakers and those working on stage and screen were launched on World Mental Health Day this year (WHO, 2019). Some individuals who are experiencing suicidal distress do search for the names of celebrities online when researching suicide methods (Biddle et al., 2012, 2018), but it is possible that details of fictional characters’ suicides may also be searched by individuals when
in distress; it is an empirical question to be addressed in future research. In addition to incorporating the names of celebrities into the SPR algorithm, it may also be appropriate to include the names of fictional characters who are particularly popular and present within the public discourse at the time.

Previous work by Arendt and Scherr (2017) suggested ways of optimizing the SPR according to the season, day of the week and proximity to family-focused public holidays and we have also made some suggestions for potential future optimization of the SPR, based on the suicide and NSSI research literature. New keywords derived from theoretical models of suicide such as the IMV model could be used to generate other phrases (e.g., “trapped” and “suicide”) to inform the SPR algorithm. Another way of developing the SPR in the future would be to refine the SPR algorithm using involvement from key stakeholders in suicide prevention, including those with lived experience of suicide ideation, suicide attempts, and bereavement after a suicide. Accordingly, Robinson et al. (2018) involved young people in the creation of materials for #ChatSafe, an initiative in Australia to help young people talk safely about suicide online. Co-development has been recently discussed as an important element within the mental healthcare “digital revolution” (Bucci et al., 2019). To this end, Google’s SPR could benefit from involving stakeholders, such as those with lived experience of suicide in order to refine the SPR algorithm further.

1.2. Ethical issues of proprietary technology in global suicide prevention efforts

In the last 50 years, we have not significantly advanced in our ability to predict suicidal thoughts, behaviours or deaths (Franklin et al., 2017). Technology and innovation have advanced the way we approach problems in health, from medical imaging techniques such as fMRI to more recent advances in dynamic blood-sugar monitoring for people with diabetes. Yet, we are not convinced that psychiatry and clinical psychology have benefited in the same way as other fields in healthcare in terms of the implementation of new technological solutions. The Experience Sampling Method (ESM; Hektner et al., 2007) and Ecological Momentary Assessment (EMA; Stone and Shiffman, 1994) offer promise in this regard, as does the use of wearable technology (Kleiman et al., 2019), as they allow the dynamic, contextual correlates of behaviour to be identified, as they occur in individuals’ daily lives (Myin-Germeyns et al., 2018). The SPR is another example of how the power and reach of a global technology, the internet, can be used to intervene in the suicidal process at points where individuals may be most vulnerable, such as at times when they are actively searching for suicide-related content online.

Whereas in the consumer technology market companies such as Apple and Samsung compete to produce more advanced smartphones and smartwatches, which drives innovation, there is little evidence of a technological arms race in suicide prevention or mental healthcare more broadly. Google, the developer of the SPR algorithm, is a highly competitive, multinational technology company and as such, algorithms including that underpinning the SPR are their proprietary intellectual capital. The SPR in essence, however, is also a public health intervention for suicide prevention on a global scale. Is it ethical for the methodology behind a potentially global suicide prevention intervention to be a trade secret? Others have also questioned the ethics of publicly listed companies, such as Google and Facebook, taking a role in suicide prevention, when their primary responsibilities as businesses are to their shareholders (Barnett and Torous, 2019). Google and Facebook are not the first to make suicide prevention a commercial entity, with companies such as LivingWorks developing and delivering safeTALK and ASIST suicide awareness and prevention training programs, respectively. Here, a key difference is that the content of these programs – that is, the intervention – is available and subject to evaluation (e.g., Bailey et al., 2017). For Google’s SPR, as with Facebook’s suicide prevention initiative, there is a lack of transparency around the content, in this case around the algorithms (Barnett and Torous, 2019).

Technologies, including Google’s SPR, are operating at the frontier of data science and public health, but also pushing the boundaries of ethics. Given the ever increasing pace of technological innovation it is paramount that technology companies and suicide prevention researchers work together in order to tackle emerging ethical issues arising from these kind of interventions, which currently operate largely beyond official regulatory oversight (Nebeker et al., 2019).

1.3. Big data, big interventions: a need for big evaluations?

The original Google blog post announcing the SPR (Zieger, 2010) states that the US National Suicide Prevention Lifeline reported receiving a 9% increase in calls following its release. No further information is provided regarding what the frame of reference was for the increase, for example, an increase in relation to the month prior to the SPR release, or an increase in relation to the same period the previous year. Additionally, no details are given explaining whether this 9% increase in calls could be directly linked to the SPR, for example by callers mentioning they had seen the Lifeline number on Google or via calls using the SPR to initiate a call directly from their web browser. To the best of our knowledge, there is no empirical evidence evaluating whether or not this SPR actually influences people’s suicide-relevant behaviour in either a positive or negative way (Stack, in press).

Whilst Arendt et al. (2019a), did an admirable job of testing the SPR in the absence of any information from Google regarding the actual algorithm, truly thorough and robust evaluation of the SPR intervention is extremely challenging without having access to the actual intervention. It is therefore imperative that Google and companies offering other similar suicide prevention interventions collaborate with independent researchers in order to evaluate these interventions. Given the potential reach of such interventions, each national suicide prevention strategy should include an action related to working with Google (or other companies) to ensure that the SPR is tailored to their country to optimize their life-saving potential.

Irrespective of access to the SPR algorithm, designing an evaluation for a global suicide prevention intervention of this kind would be methodologically and logistically challenging. A major review of suicide prevention strategies, conducted by Zalsman et al. (2016) highlighted the lack of “… gold standard … randomised trials” (p. 10) evaluating the efficacy of suicide prevention interventions. Zalsman and colleagues also state that “potential interventions using new social media, mobile technologies, and continuous monitoring of large data sets seem to be the next field to explore in the coming decade” (p. 10). Large-scale, internet-based interventions like the SPR are not suited to evaluation using traditional gold-standard methods of randomised controlled trials (RCT), as is the case for other types of universal, population-level interventions for preventing suicide, such as barriers on bridges and restrictions on particular types of medications (Hawton and Pirkis, 2017). There is a growing understanding within the field of prevention science that some interventions and contexts are ill-suited to evaluation via RCT and that alternatives are necessary (Henry et al., 2017). To assign only some individuals, at random, to receive the SPR when searching for suicide information online would be unethical. Following Hawton and Pirkis (2017) suggestions regarding evaluation of universal interventions, given that the SPR is currently only available in 14 countries an alternative design for evaluating the SPR may be a roll-out or wait-list design. Here, all individuals would eventually receive the intervention, eliminating the ethical issue of randomizing participants to a potentially life-saving intervention, but comparisons would be made before and after the intervention was introduced. In the case of the SPR, crisis line calls or crisis line link visits could be assessed prior to roll-out of the SPR in a particular region, then again post roll-out, as could the actual population suicide rates. In an Australian study evaluating the “Man Up” suicide prevention initiative website, King et al. (2019) used Google Analytics to analyse the number of website
visitors and how often they clicked on the links to support services given on the website. This provided some indication of whether or not presenting sources of potential support translates into individuals accessing them. Evaluation of this kind should be performed as a matter of urgency for the SPR. The use of Google analytics tools, such as Google Trends, may nonetheless be the optimal method of evaluating the SPR. A study by Tran et al. (2017) demonstrated that Google Trends data has low validity and reliability for forecasting actual suicidal behaviour as measured by national suicide rates. This study also raises further questions regarding what appears to be an underlying assumption of the SPR – that online searches for suicide-related content do indeed relate to actual suicidal behaviour and that this relationship may be disrupted by displaying the SPR. Tran et al.’s (2017) study yet again highlights the need for interventions to have a solid basis in evidence, not just in terms of their “active ingredients”, but also in terms of their rationale.

Recent times have seen a proliferation of apps and other technological “solutions” to address mental health problems, including suicide. Nonetheless, a recent review contended that there was insufficient evidence to conclude that any available mental health app could be considered as evidence-based (Lui et al., 2017). It is crucial, therefore, for individuals to be aware that simply because a mental health app or online tool, such as the SPR, exists does not mean that it is evidence-based or that it has been subject to rigorous scientific evaluation. Furthermore, the rate of technological development has clearly outpaced evaluation (Bucci et al., 2019), which is especially concerning for a universal suicide prevention initiative such as the SPR, where potential harms or benefits – both of which are currently unknown – are occurring on a global scale.

2. Conclusions

That a major technology company like Google has moved into developing interventions for suicide prevention sends a clear message that suicide prevention has yet to benefit from a technological arms race as witnessed in other healthcare fields, Google’s involvement and the potential reach of the SPR intervention brings opportunities. These advances, however, are also accompanied by critical ethical and methodological considerations, some of which do not yet have optimal solutions. Arendt et al. (2019a), study has pinpointed areas where the SPR can and should be optimized, specifically by displaying it when individuals search for details of celebrity suicides; unfortunately too often a source of details regarding suicide methods (e.g., Biddle et al., 2018). These optimizations can and should be evidence-based, by drawing upon theoretical frameworks from suicide, such as the IMV model (O’Connor and Kirtley, 2018) and NSSI research (e.g., the Cognitive Emotional Model; see Hasking et al., 2017). Interventions to prevent suicide should be evidence-based and the SPR is no exception to this, however, the ability of independent researchers to evaluate the SPR is severely hampered by the fact that the algorithm underlying the SPR is proprietary. Furthermore, it is unclear whether or not the behavioural pathway that the SPR attempts to disrupt (searching for suicide-related content relates to suicidal behaviour) is more than an assumption, as recent work indicates low reliability and validity of Google Trends data for forecasting national suicide rates (Tran et al., 2017).

Transparency regarding such algorithms is crucial (Barnett and Torous, 2019), and remains a key ethical issue to be tackled by technology companies in collaboration with researchers. Technologies that purport to have a clinical benefit should first pass a threshold of demonstrating their value, prior to their release (Nebeker et al., 2019). There is currently no empirical evidence regarding the SPR’s benefits or harms. As suicide prevention increasingly becomes everyone’s business, so too must the rigor in developing and evaluating prevention initiatives.

Acknowledgments

Olivia Kirtley is supported by a postdoctoral fellowship from an FWO Odysseus grant to Prof. Inez Myin-Germeys, KU Leuven (Fonds Wetenschappelijk Onderzoek - Vlaanderen, GOF8416N).

References


