

## Review



## Practices for monitoring and responding to incoming data on self-injurious thoughts and behaviors in intensive longitudinal studies: A systematic review

Kate H. Bentley<sup>a,b,\*</sup>, Joseph S. Maimone<sup>a,c</sup>, Erin N. Kilbury<sup>c,a</sup>, Marshall S. Tate<sup>c</sup>, Hannah Wisniewski<sup>d</sup>, M. Taylor Levine<sup>a</sup>, Regina Roberg<sup>a</sup>, John B. Torous<sup>d</sup>, Matthew K. Nock<sup>a,c</sup>, Evan M. Kleiman<sup>e</sup>

<sup>a</sup> Massachusetts General Hospital, Department of Psychiatry, United States of America

<sup>b</sup> Harvard Medical School, Department of Psychiatry, United States of America

<sup>c</sup> Harvard University, Department of Psychology, United States of America

<sup>d</sup> Beth Israel Deaconess Medical Center/Harvard Medical School, Department of Psychiatry and Division of Digital Psychiatry, United States of America

<sup>e</sup> Rutgers University, Department of Psychology, United States of America

## ARTICLE INFO

**Keywords:**  
Suicide  
Self-injury  
Ecological momentary assessment  
Mobile health

## ABSTRACT

Advancements in the understanding and prevention of self-injurious thoughts and behaviors (SITBs) are urgently needed. Intensive longitudinal data collection methods—such as ecological momentary assessment—capture fine-grained, “real-world” information about SITBs as they occur and thus have the potential to narrow this gap. However, collecting real-time data on SITBs presents complex ethical and practical considerations, including about whether and how to monitor and respond to incoming information about SITBs from suicidal or self-injuring individuals during the study. We conducted a systematic review of protocols for monitoring and responding to incoming data in previous and ongoing intensive longitudinal studies of SITBs. Across the 61 included unique studies/samples, there was no clear most common approach to managing these ethical and safety considerations. For example, studies were fairly evenly split between either using automated notifications triggered by specific survey responses (e.g., indicating current suicide risk) or monitoring and intervening upon (generally with a phone-based risk assessment) incoming responses (36%), using both automated notifications and monitoring/intervening (35%), or neither using automated notifications nor monitoring/intervening (29%). Certain study characteristics appeared to influence the safety practices used. Future research that systematically evaluates optimal, feasible strategies for managing risk in real-time monitoring research on SITBs is needed.

Self-injurious thoughts and behaviors (SITBs), including those that are suicidal and nonsuicidal in nature, are major public health problems. More than 47,000 Americans die by suicide each year, which works out to about 130 suicide deaths per day (Centers for Disease Control and Prevention, 2020). Alarming, rates of death by suicide increased year-by-year in 17 out of the past 20 years (Centers for Disease Control and Prevention, 2020). Suicidal thoughts and nonfatal suicidal behaviors (e.g., suicide plans, preparatory actions) confer risk of future suicide (Ribeiro et al., 2016) and have significant public health costs in their own right (Shepard, Gurewich, Lwin, Reed, & Silverman, 2016). Nonsuicidal self-injury (NSSI) is also quite common, especially among young people, with estimated prevalence rates ranging from 5.5 to 17.2%

across age groups (Swannell, Martin, Page, Hasking, & St John, 2014). NSSI is among the strongest known predictors of future suicidal behavior (Ribeiro et al., 2016) and has clinically important consequences such as worsened psychopathology, physical injury, and lower overall functioning (e.g., Briere & Gil, 1998; Nock, Joiner Jr, Gordon, Lloyd-richardson, & Prinstein, 2006; Turner, Austin, & Chapman, 2014).

Despite decades of research on these problems, we are not yet able to accurately predict and prevent the onset or escalation of suicidal and nonsuicidal SITBs. One reason for this lack of progress is that we still have poor understanding of many of the basic properties of SITBs, as well as the time-varying, proximal risk and protective factors that contribute to their occurrence (Franklin et al., 2017; Glenn & Nock,

\* Corresponding author at: Massachusetts General Hospital, Department of Psychiatry, 1 Bowdoin Square, 6th Floor, Boston, MA 02114, United States of America.  
E-mail address: [kbentley@mgh.harvard.edu](mailto:kbentley@mgh.harvard.edu) (K.H. Bentley).

2014). This may be in part because, unlike many other psychological phenomena, it is both practically and ethically challenging to gather information on SITBs via direct observation in a research laboratory or office. Thus, most previous studies have used traditional assessment methods that involve asking suicidal or self-injurious individuals to report the number of times they experienced or the intensity of SITBs over the past week(s), month(s), or year(s). Though not without value, this strategy is associated with retrospective recall biases, is subject to demand characteristics, and may not shed light on when and why SITBs occur across shorter, more clinically relevant time windows (e.g., minutes, hours, days).

Fortunately, recent advances in mobile technologies such as smartphones and wearable biosensors have made it possible to study SITBs in real-time as they naturally occur (Kleiman & Nock, 2017). Although such technologies have existed for decades, they have become exponentially more accessible to researchers in recent years (as for example, the vast majority of research participants now already own a smartphone; Pew Research Center, 2019) and thus scalable across increasingly large samples: an especially important consideration for studies of SITBs, as relatively low-base rate phenomena. Using smartphones and biosensors to gather real-time data on SITBs (and factors thought to contribute to SITBs, such as negative affect, social activity, and sleep) has the potential to not only advance our understanding of when and why SITBs occur, but also inform the development of just-in-time interventions that provide support during periods of risk (Coppersmith et al., 2020; Nahum-Shani et al., 2018). The promise of digital monitoring technologies for SITBs has been widely recognized, as evidenced in part by the recent proliferation of published studies using ecological momentary assessment (EMA) or similar methods to study SITBs (for reviews see: Gee, Han, Benassi, & Batterham, 2020; Hepp et al., 2020; Kleiman & Nock, 2017; Rabasco & Sheehan, 2021; Rodríguez-Blanco, Carballo, & Baca-García, 2018; Sedano-Capdevila, Porras-Segovia, Bello, Baca-García, & Barrigon, 2021).

Collecting information about SITBs in real-time, however, poses important safety, ethical, and methodological concerns. Although recent findings show that asking participants frequently and repeatedly about suicidal thoughts in EMA studies is not iatrogenic (e.g., Coppersmith et al., 2020; Rabasco & Sheehan, 2021), there are not yet well-established best practices on key issues such as whether to monitor and how to respond to incoming data on current SITBs (e.g., Berman & Carter, 2020; Nock et al., 2020). For example, when participants submit a survey response indicating current suicidal intent that researchers can access in real-time (or very close to it), should the study team intervene? How should the study team determine when an intervention is needed? What should the intervention involve? Should whether and how researchers monitor and respond to incoming data on SITBs depend on the study population, the monitoring period (e.g., during hospitalization versus after discharge), or other factors (e.g., whether participants have a provider, whether participation is anonymous)? Moreover, as these studies continue to be scaled, what interventions are feasible and sufficiently low-cost to deploy across bigger, national and possibly even international samples? Critically, what is the impact of monitoring and intervening on participants reporting the very suicide-related outcomes we most need to study, and thus ultimately better understand and prevent? These are only a few of many questions that researchers doing real-time monitoring research on SITBs grapple with. Such issues are also highly pertinent to other key members of the research process, including the funding agencies supporting this rapidly growing area of research, Institutional Review Boards tasked with ensuring that participant safety is prioritized, and study participants themselves (or their parents), as well as the clinicians who may be treating the patients participating in these studies.

Fortunately, a few recent investigations have begun to address some of these questions. Our team organized a panel of 24 experts from psychology, psychiatry, law, bioethics/IRB, funding agencies, and lived experience who issued a consensus statement (using a modified Delphi

process) on the ethical and safety practices for conducting real-time monitoring studies of individuals at risk for suicide and related behaviors (Nock et al., 2020). Consensus was reached about some, but not all, issues addressed. Regarding monitoring of incoming data, for example, there was strong (about 90%) consensus that incoming data on suicide risk should be reviewed at least every weekday (or in real-time, when the data collection platform permits this), whereas it was less clear how to determine level of current risk based on EMA data. Regarding corresponding interventions, there was also strong (about 94%) consensus that when participants provide a “high-risk” response, the study team should reach out to them directly to conduct a suicide risk assessment as soon as possible. For “low/moderate risk” responses, however, there was less (about 61%) agreement on whether an automated pop-up message within the survey (e.g., a screen on the app encouraging help-seeking or providing a safety plan) is sufficient. Others (e.g., Schatten et al., 2020) have also recently offered recommendations for potential best practices in suicide risk assessment and intervention in clinical research, including EMA studies.

These efforts, however, have not systematically summarized the specific safety practices that researchers conducting intensive longitudinal studies of suicidal or self-injuring individuals have used to date or are currently using. A recent report (Bai, Babeva, Kim, & Asarnow, 2020) focused on the risk management strategies in prior EMA studies of youth with SITBs; however, it did not include studies of adults (which comprise the majority of EMA SITB research to date) and reviewed only six articles (most of which were published four or more years ago), rendering it difficult to draw conclusions about factors that may impact safety protocols. A systematic, comprehensive summary of practices for monitoring and responding to incoming data on SITBs in recent or ongoing digital monitoring studies would offer scientists and other stakeholders current information about the most commonly used practices, and how these may vary depending on key study characteristics such as the design or population used—thus potentially allowing researchers to align their studies with a common standard. Accordingly, we set out to conduct the first systematic review of practices for monitoring and responding to incoming intensive longitudinal data used in both previously published and ongoing (grant-funded) digital monitoring studies of SITBs across populations. We were especially interested in identifying whether safety protocols varied as a function of sample size, type (e.g., clinical versus community-based participants), and age (adolescents versus adults), or duration of the monitoring period, as well as trends over time.

## 1. Method

This systematic literature review complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (PRISMA; Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009).

### 1.1. Search strategy and selection

#### 1.1.1. Published articles

We conducted a comprehensive literature search of PsycINFO, PubMed, Embase, Cochrane, and Web of Science from database inception through April 2021 to identify studies eligible for this systematic review. Terms (including alternates) related to digital monitoring or EMA (*android, app, application, cellular, computer program, daily diary, device, ecological momentary, EMA, experience sampling, ESM, handheld, iPhone, iPad, iOS, mobile, platform, phone, short message service, SMS, software, smartphone, tablet, technology, text message*) were entered simultaneously with suicide or self-injury related terms (*automutilation, cutting, deliberate self-harm, DSH, NSSI, parasuicide, self-harm, self-injury, self-injurious behavior, self-mutilation, suicide*). We also searched the reference sections of identified review articles (e.g., Rodríguez-Blanco et al., 2018; Sedano-Capdevila et al., 2021) for other publications.

Fig. 1 presents a PRISMA diagram of the study selection process. A total of 1985 unique publications were identified by initial literature searches and the reference sections of review articles; based on reviewing each title and abstract, 143 of these were deemed eligible for further review. After reviewing full texts, 100 publications met criteria for inclusion. Publications that reported on two or more distinct samples (e.g., “study 1” of psychiatric inpatients and “study 2” of community participants) were treated as separate studies for purposes of this review because safety practices could differ between the two samples. Of the 100 unique publications, 82 (82.0%) either reported information on data monitoring and response protocols in the article or the author(s)

responded to our request to provide us with this information (via a structured data extraction form). As some researchers reported on the same sample/study (and thus used the same safety protocols) in multiple publications, there were a total of 43 unique published studies included in our review.

1.1.2. Grants

To identify current grants, we searched the National Institute of Health Research Portfolio Online Reporting Tools (NIH RePORTer) database for active or completed projects (through April 2021). Terms related to digital monitoring or EMA (e.g., *smartphone*, *EMA*, *ecological*

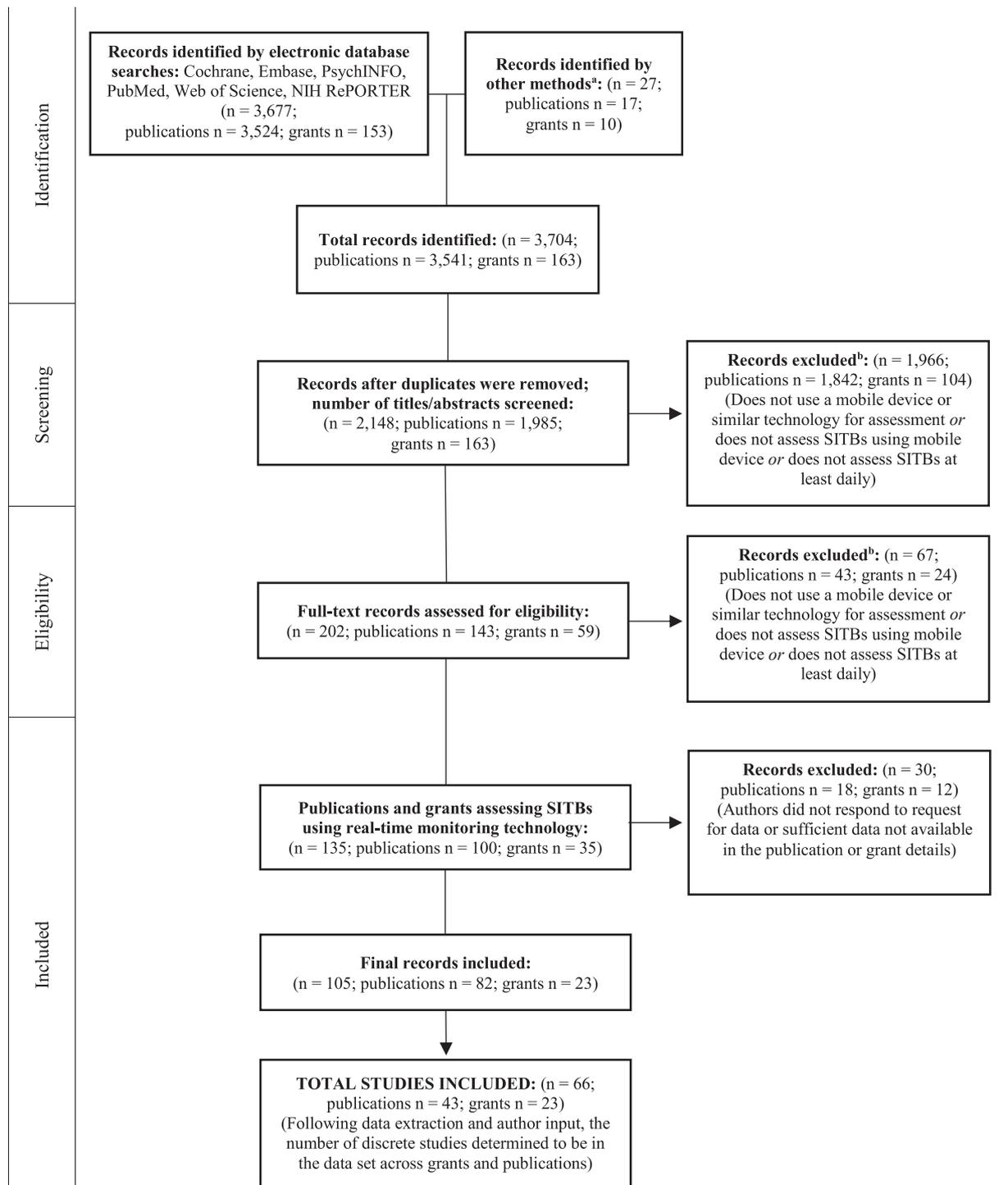


Fig. 1. Prisma flow diagram.

momentary assessment, app) were entered simultaneously with suicide or self-injury related terms (e.g., *suicide*, *self-harm*, *NSSI*, *suicidal*). We also manually reviewed all active grants funded by the American Foundation for Suicide Prevention (AFSP; <https://afsp.org/research-grants/>) for projects using digital monitoring or EMA methods as of the 2020 award cycle. This process resulted in a total of 35 unique ongoing grant-funded studies eligible for inclusion in this review. We emailed the Principal Investigators (PIs) for eligible studies requesting information (via structured data extraction form) on their safety protocols; of these, information was obtained for 23 unique studies not already captured in the publications.

### 1.2. Inclusion and exclusion criteria

To be included in this systematic review, studies had to first use a mobile device or similar technology to assess variables of interest over time. We did not require use of a smartphone; studies that used computers or tablets were included as long as other inclusion criteria were met. Second, included studies had to assess suicidal or nonsuicidal SITBs using the technology identified in the first criterion. Third, studies were required to assess SITBs at least daily; if, for example, studies used EMA to capture other phenomena (e.g., negative affect, depressive symptoms, emotion dysregulation), but did not explicitly report using EMA or similar data collection methods to repeatedly assess SITBs, then the study was not eligible. Third, studies had to either provide information on their protocol for monitoring or responding to incoming data in the publication (for articles) or via email correspondence with the author(s) or grantee(s). Finally, studies had to appear in a publication (or be supported by an active NIH or AFSP grant as described above) and present original data not already reported in another study.

### 1.3. Data extraction

Relevant information was systematically extracted by two independent coders for each publication. Discrepancies were resolved by discussion; when a discrepancy couldn't be easily resolved by the two independent coders, the first author made a final decision. For articles that did not report information on safety practices and for all grants, information was requested from the research team via email using a standardized data entry form; two contact attempts were made via email. Recorded information included (a) sample characteristics (N, age [e.g., adolescents versus adults], type [e.g., outpatients, inpatients, community, students], primary problem/diagnosis); (b) study design characteristics (design [assessment-only versus assessment plus intervention], duration of monitoring period, types of SITBs assessed, device used to collect data on SITBs); and (c) protocols for monitoring and responding to incoming survey data.

The following variables (with binary or open-ended responses) summarized information on protocols for monitoring and responding to incoming survey data: (c1) did the EMA or daily diary survey display a safety-related automated pop-up notification; (c2) what was the threshold for triggering the pop-up; (c3) what was the content of the pop-up (e.g., provide hotline numbers, recommend emergency services, etc.); (c4) were incoming data monitored for "high-risk" responses (either via an automated alert system or manual study staff review) during the study; (c5) how frequently were data reviewed, (c6) what was the threshold for defining a response as "high-risk," (c7) what was the protocol for responding to "high-risk" responses, and (c8) what study staff member (e.g., research assistant, study clinician/PI) was responsible for responding.

## 2. Results

### 2.1. Description of included studies

Characteristics of the 66 unique included studies (publications and

grants) are shown in **Supplemental Tables 1 and 2**. For all published articles, years of publication spanned 2007 to 2021, with well over half (69.5%) being published within the past four years (in or after 2017).

#### 2.1.1. Sample characteristics

Across all studies, there were a total of 5918 participants (3513 in published studies and 2405 targeted for grant-funded studies). The mean number of participants per published study was 81.7 ( $SD = 85.3$ , range = 4–491). The mean number of targeted participants per ongoing grant-funded study was 114.5 ( $SD = 55.8$ , range = 30–205). The majority (69.7%) of studies (published and grant-funded) used exclusively adult or young adult samples (minimum age 18); another 19.7% used exclusively adolescent samples, 7.6% used an adolescent/young adult sample (e.g., 15- to 25-year-olds), and two studies enrolled adolescents and the full age range of adults. Of the studies that provided information on sample type, 28.1% used outpatient, 26.6% community-based, 20.3% inpatient, 9.4% student, and 4.7% used other hospital-based (e.g., emergency room) samples; 10.9% used a combination (e.g., community and student). Of the studies that provided such information, 69.8% required that participants report current or a history of SITBs for inclusion in the study, whereas 30.2% of studies did not.

#### 2.1.2. Study design characteristics

For studies with a fixed EMA or daily diary period (93.9%), the mean duration of the monitoring protocol was 33.2 days ( $SD = 40.9$ , range = 4–252). For 6.1% of studies, the duration of the monitoring protocol varied across participants as it corresponded to the length of hospital stay; of these, two studies (ongoing grants) also extended the monitoring period to 28 days after hospital discharge. The vast majority of studies (81.8%) was assessment-only, with 18.2% of studies classified as assessment plus intervention. Studies most commonly collected intensive longitudinal data on both suicidal and nonsuicidal SITBs (37.9%) or only suicidal thoughts or behaviors (34.8%), whereas 27.3% collected data only on NSSI. Most (72.7%) studies specified using a smartphone (app) to collect intensive longitudinal data on SITBs, whereas 16.7% used other devices (e.g., palm pilots, electronic tablets), 9.1% a website (that may have been accessed via mobile device or computer), and one study used email.

### 2.2. Protocols for monitoring and responding to incoming data on SITBs

#### 2.2.1. Automated pop-up notification

Of the studies that provided information about the use of an automated safety-related pop-up notification, just under half (47.6%) reported using a pop-up in the EMA or daily diary survey whereas 52.4% did not use a pop-up. The following thresholds were used to trigger the pop-up (some studies used more than one threshold so percentages add to over 100%): presence of any current or recent (e.g., past 12 or 24 h) suicidal ideation (33.3%), severity of current or recent suicidal ideation (33.3%), presence of recent suicide attempt (21.2%), severity of current or recent suicidal intent (16.7%), presence of current or recent suicide plan (16.7%), presence of recent NSSI (16.7%), presence of any current or recent suicidal intent (13.3%), severity of recent NSSI (3.3%), and other (e.g., homicidal ideation; 6.7% studies).

Of the studies that did not use a pop-up, 30.0% indicated that they provided similar resources as would be shown in a pop-up, but in an end-of-survey page that was always displayed (i.e., not contingent on a response threshold). Content of pop-ups or end-of-survey messages included encouragement to call a crisis line or similar (65.0%), contact existing supports (e.g., therapist, family; 55.0% studies), use emergency services (e.g., call 911, go to the emergency room; 47.5% studies), contact the study team (14.6% studies), or use a safety plan (10.0% studies); 15.0% of studies also reported including a reminder that data are not monitored in real-time in the pop-up or end-of-survey message.

Regarding whether pop-up use varied as a function of study characteristics, over half (55.3%) of studies conducted in the past four years

used a pop-up whereas the minority (25.0%) of studies conducted before 2017 used a pop-up. The vast majority (88.9%) of studies with a monitoring period of at least 30 days – and the minority (29.3%) of studies with less than 30 days of monitoring – used a pop-up. Whereas the majority (59.5%) of studies in clinical samples used a pop-up, the minority (33.3%) of studies in non-clinical samples used a pop-up. In studies with at least 100 participants, it was somewhat more common to use a pop-up versus not (59.1% and 40.9%), whereas in studies of under 100 participants, it was slightly less common to use a pop-up versus not (43.6% and 56.4%). It was slightly more common for adolescent studies to use a pop-up (57.9%), whereas it was slightly less common for adult studies to use a pop-up (43.2%).

### 2.2.2. Monitoring incoming data on SITBs

Over half (60.6%) of studies monitored incoming data on SITBs during the study, whereas 39.4% of studies did not. It was most common for studies to either monitor incoming data or use an automated pop-up message but not both (36.4% of studies did one or the other, whereas 34.8% of studies did both, and 28.8% of studies did neither).

For studies in which incoming data on SITBs were monitored, the following thresholds were used to define survey responses as “high-risk” (some studies used more than one so percentages add to over 100%): severity of current or recent (e.g., past 12 or 24 h) suicidal intent (36.3%), presence of current or recent suicide plan (30.3%), severity of current or recent suicidal ideation (27.3%), presence of recent suicide attempt (27.3%), presence of any current or recent suicidal intent (21.2%), presence of recent NSSI (18.2%), presence of any current or recent suicidal ideation (12.1%), and other (e.g., homicidal ideation, low ability to resist suicidal urges; 9.1% studies). Data were reviewed at the following frequencies: real-time (e.g., via automated alerts; 57.9%), once daily (28.9%), more than once daily (e.g., 7.9%), and less than daily (5.3%).

In terms of whether monitoring incoming data varied as a function of study characteristics, 65.3% of studies from the past four years monitored incoming data whereas just under half (47.1%) of studies before 2017 monitored incoming data. The vast majority (78.9%) of studies with monitoring periods of at least 30 days and just over half (55.8%) of studies with monitoring periods of under 30 days monitored data. The majority (75.0%) of studies on adolescents and just over half (54.3%) of studies on adults monitored incoming data. The majority of studies with samples of at least 100 participants (64.0%) and less than 100 participants (59.0%), and the majority of studies using clinical (64.1%) and nonclinical samples (60.0%), monitored incoming data.

### 2.2.3. Responding to incoming data on SITBs

Of the studies that monitored incoming data on SITBs throughout the study, 94.7% of studies reached out to the participant (or for some adolescent studies, the participant’s parents first, followed by the participant) directly (most often by phone) to conduct a suicide risk assessment. For all studies that provided such information, the risk assessment was conducted by a more senior member of study staff than a research assistant (e.g., PI, p.H.D.-level staff member, or clinician). For 26.3% of studies, a (non-study) clinical provider was notified (in addition to a risk assessment or as a standalone response); all of these studies were conducted in clinical samples. Finally, 7.9% of studies reported systematically providing information about referrals to the participant as part of the high-risk response protocol.

### 2.2.4. Other common safety practices

There were several other relevant safety practices that authors or PIs described via open-ended responses in our data extraction form. Because we did not systematically assess these practices for all included studies, we do not report percentages here but only note other common practices endorsed. A number of studies emphasized to participants during the informed consent process that EMA or daily diary data are not monitored in real-time. Others described giving participants clinical or crisis

resources at the start of the study, which may have included a study-specific phone number for emergencies (e.g., for an on-call study clinician). Finally, other studies reported developing a personalized safety plan with participants at the start of the study, which in some cases could be accessed in the smartphone app used for EMA.

## 3. Discussion

Intensive longitudinal data collection methods have great potential to increase the understanding of how SITBs unfold in real-time, improve short-term prediction of these clinically severe and sometimes fatal outcomes, and inform the development of much-needed interventions. Researchers conducting these studies, however, face complex ethical and practical considerations, namely about whether and how to respond to incoming real-time data on current risk for suicide and related behaviors. This review of safety protocols for monitoring and responding to incoming intensive longitudinal data on SITBs revealed three main findings. First, the number of intensive longitudinal studies of SITBs has recently proliferated, with well over half of all studies published in the past four years. Second, there was no single most common approach to handling incoming real-time data on SITBs, with roughly half of studies using automated safety-related pop-up notifications and roughly half of studies not using pop-ups, and about 60% of studies monitoring and taking action in response to “high-risk” responses and about 40% of studies not monitoring or responding to incoming data. Third, study recency, population, and duration appeared to affect whether these safety practice(s) were used. These results have important implications for scientists and other stakeholders involved in this rapidly growing area of suicide and self-injury research.

The current finding that studies were fairly evenly split between using and not using an automated pop-up is in line with our team’s recent report of only about 60% consensus among a panel of experts that a pop-up notification should be triggered by non-imminent risk responses (Nock et al., 2020). In terms of why researchers may decide not to use an automated pop-up notification containing safety resources, we speculate that there could be concern about the potential for habituation, for example if displayed too frequently or consistently. It is also likely that the limitations of earlier data collection platforms influenced results regarding pop-up use, as over half of publications from the past four years used a pop-up, compared with roughly 25% of earlier publications. For example, the ability to have conditional logic (e.g., display a screen based on a response to a prior item) has been an advanced feature available only in the past few years. Studies with longer monitoring periods and in clinical samples were also more likely to use a pop-up (89% and 60%, respectively) than those of shorter duration and in nonclinical samples (29% and 33%, respectively). Of note, the most common threshold used to trigger a pop-up was self-reported presence or severity of current or recent (e.g., past 12 or 24 h) suicidal ideation; however, studies were highly variable in the thresholds (e.g., suicidal intent, suicide plan, NSSI, etc.) used to trigger both pop-ups and more intensive interventions (discussed below) overall.

That a substantial proportion (about 40%) of studies did not monitor or respond to incoming data on SITBs represents a departure from our team’s recent report of very strong (roughly 90% or more) consensus among experts that data should be reviewed at least daily, and in the case of “high-risk” responses, the study team should reach out to the participant for a risk assessment as soon as possible (Nock et al., 2020). This divergence may at least partially reflect the limitations of technologies used in earlier studies that did not allow data to be uploaded remotely for the study team to access in an ongoing way; indeed, whereas about 65% of studies from the past four years monitored and responded to incoming data on SITBs, less than half (about 47%) of studies conducted before 2017 monitored incoming data.

There are also other potentially compelling reasons why researchers may decide not to monitor incoming intensive longitudinal data on SITBs. First, studies that enroll participants anonymously (e.g., via

online recruitment methods) are necessarily limited in the potential to intervene (at least via human outreach) to incoming data. Second, in any study where psychological phenomena are assessed repeatedly or participants know they are being monitored, reactivity can be a concern (e.g., Reynolds, Robles, & Repetti, 2016; van Ballegooijen et al., 2016). For studies in which participants are aware the study team will not only monitor their data, but also take action if they provide a “high-risk” response (which in this review, was most often defined as severity of suicidal intent), reactivity may be especially relevant, as participants could underreport SITBs (or stop responding to study surveys entirely) to avoid an unwanted intervention. Participants may, for example, perceive the study team contacting them for a suicide risk assessment as unhelpful or burdensome; this may be of particular concern to adolescents, for whom the protocol may also involve contacting their parents. Participants may also be concerned that the police will ultimately be notified, which could result in serious negative (even life-threatening) consequences, particularly for Black and Hispanic participants (Buehler, 2017; Edwards, Lee, & Esposito, 2019); indeed, some PIs informed us that they no longer give participants information about traditional emergency services for this very reason. It is also plausible that for some participants, being called for a risk assessment may reinforce SITB reporting.

If monitoring and responding to incoming data on SITBs does influence responding behavior, this could muddle the validity of resultant study data, potentially limiting researchers’ ability to observe with precision (and thus learn the most about) the phenomena of primary interest in these studies. To our knowledge, however, these concerns about monitoring and responding to incoming data in intensive longitudinal studies of SITBs are currently largely speculative or based in anecdotal reports. Thus, future studies that randomly assign participants to various “data will be monitored” or “data will be not monitored” (or “pop-up” versus “no pop-up”, etc.) conditions, or use within-person experimental methods such as single-case experimental design (Barlow, Nock, & Hersen, 2009; Bentley, Kleiman, Elliott, Huffman, & Nock, 2019) to systematically vary the impact of different monitoring approaches on participants’ self-reports of suicidal thoughts over time would be valuable. The newer micro-randomized trial (Klasnja et al., 2015) may also be leveraged to rigorously evaluate the effects of intervening (and specific types of interventions; e.g., pop-up notification, human or newer automated forms of risk assessment interventions, as discussed below) on responding behavior. As noted above, individual-level factors must also be taken into account in future research in this area, as responses may vary between participants. A critical line of work, of course, must determine *when* people are at highest risk and require an intervention, as self-reported suicidal thoughts do not necessarily indicate that suicidal behavior is imminent (Nock et al., 2020); indeed, improving the prediction of proximal suicide risk is a main aim of this overall body of research.

Another key consideration is the feasibility of monitoring and responding to incoming data, as these approaches tend to require considerable staff, at least as traditionally deployed. For example, our research team has a rotating doctoral-level staff member on-call (seven days a week) who, upon receiving automated, real-time alerts that a “high-risk” response has been submitted, calls the participant (or the parent if the participant is an adolescent) directly for a semi-structured risk assessment. Indeed, in this review, we found that greater percentages of both larger (at least 100-participant) and longer (at least 30-day) studies monitored and responded to incoming data than smaller and shorter studies. This may reflect a tendency for the larger and longer studies to be better resourced (e.g., federally funded) and thus more realistically able to implement these safety practices. Especially as these studies continue to be scaled, which is much needed in order to capture low base rate phenomena such as suicide attempts and deaths, relying on a trained clinician or similar to conduct such risk assessments becomes less feasible. Recent advances in mobile technologies have the potential to facilitate *automated*, potentially highly efficient risk assessment

strategies (i.e., that do not require a human calling) and deployment of specific types of notifications/alerts either delivered directly to the participant (e.g., crisis line resources, app-based coping skills) or existing supports (e.g., clinical providers, significant others, etc.) (National Institute of Mental Health, n.d.). These strategies could also be leveraged in anonymous studies. Another relevant line of future work is to systematically compare the risk of adverse events associated with certain intervention strategies versus others; it is possible that some interventions (e.g., sending help directly to a participant) may have higher risk of resulting in an adverse event, especially for certain groups. Prioritizing participant safety does not consist only of protecting against risk of immediate suicidal behavior, but also other potentially harmful outcomes.

Ultimately, researchers must balance the feasibility and potential risks of undesirable effects on data integrity or quality associated with monitoring and responding to incoming data with what may be an ethical obligation to take action when it is learned that a participant may be at imminent risk of ending their life. Although monitoring and responding to incoming data indicating high current levels of suicide risk and triggering “pop-up” messages with crisis resources are currently commonly used strategies to help meet this ethical obligation, other (non-mutually exclusive) strategies include emphasizing (and gauging participants’ understanding of) whether or not data will be monitored during an informed consent process, providing local and general crisis resources at the start of the study, ensuring participants have access to a provider for emergencies, and developing personalized safety plans. Researchers may also consider factors such as sample age and type when weighing this balance. For example, here we observed that approximately 75% of adolescent studies monitored and responded to incoming data, compared to just over half (54%) of adult studies. This trend is in line with our recent report indicating strong consensus among experts that human outreach after a high-risk response is preferred for minors. A related area for future work is to systematically solicit and summarize the perspectives of participants engaged in real-time monitoring of SITBs regarding their preferred safety practices in order to fully represent the full range of stakeholders (e.g., Kleiman et al., 2021).

### 3.1. Limitations

There are a number of limitations to the current review that must be mentioned. For one, we did not systematically assess all possible potential safety-related practices used in each study, which limited our focus primarily to the use and nature of automated safety-related pop-up notifications and monitoring/responding to incoming data. Second, we only included studies for which the authors or PIs responded to our inquiry (or there was sufficient information provided in the publication on safety practices), which may have biased our final pool of studies (possibly toward those that *did* include the primary safety practices of interest); along these lines, we did not include a structured assessment of study bias. Additionally, by only searching for ongoing grants funded by the NIH or AFSP, we did not include grants funded by other sources. Last, just as in any review, it is possible that we missed eligible publications and ongoing grants in our search.

### 3.2. Conclusions

Despite these limitations, this review had key strengths, including a comprehensive, systematic summary of published literature in a rapidly proliferating area of research, consideration of key measures taken to prioritize safety in intensive longitudinal studies of SITBs, and attention to study-level factors that may help determine what safety practices to employ. It is our hope that this review serves as a useful resource for researchers and other key stakeholders tasked with attending to the complex safety and ethical considerations in intensive longitudinal studies of individuals at risk for suicide. Many directions for future work remain, including developing and testing scalable approaches to

managing risk in real-time monitoring studies of suicidal and related behaviors, and determining what suicide risk reduction strategies are both efficient and effective given the current internal and external context for any individual.

### Role of funding sources

This work was supported by funding from the National Institute of Mental Health (K23MH120436, PI: Bentley; U01MH116928, PI: Nock). The NIMH had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the manuscript for publication.

### Contributors

Kate Bentley, John Torous, Matthew Nock, and Evan Kleiman designed the study, wrote the protocol and conducted outreach to authors of included studies. Joseph Maimone, Erin Kilbury, Marshall Tate, Hannah Wisniewski, Meghan Levine, and Regina Roberg conducted the systematic review and data extraction of literature. Kate Bentley conducted the statistical analysis and wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

### Declaration of Competing Interest

Dr. Nock receives textbook royalties from Macmillan and Pearson publishers and has been a paid consultant in the past year for Microsoft and for a legal case regarding a death by suicide. He is an unpaid scientific advisor for TalkLife and Empatica. No other authors have any conflicts of interest to report.

### Acknowledgements

We would like to thank all of the researchers who responded to our requests for information on the safety protocols included in their studies. Specifically, we would like to thank Linda J. Alpert-Gillis, Holly E. Andrewes, Michael D. Anestis, Michael F. Army, Randy Auerbach, Enrique Baca-Garcia, Jeffrey S. Ball, Todd M. Bishop, Peter Bowen, Konrad Bresin, Laia Briones-Buixassa, Taylor A. Burke, Nicole M. Cain, Andrew M. Chanen, Laurence Claes, Philippe Courtet, Eimear Crowe, Ewa K. Cysz, Daniel P. Dickstein, Kirsten H. Dillon, Alexandre Y. Dombrowski, Tory A. Eisenlohr-Moul, William D. Ellison, Scott G. Engel, Skye Fitzpatrick, William Fleeson, Thomas Forkmann, Heidi Glaesmer, Catherine R. Glenn, Tina R. Goldstein, Fanny Gollier Briant, Kathryn H. Gordon, Ilana Gratch, Nina Hallensleben, Jessica L. Hamilton, Jill M. Hooley, Sarah Horowitz, Marlies Houben, Benjamin N. Johnson, Michael Kaess, Glenn Kiekens, Nathan A. Kimbrel, Cheryl A. King, David E. Klonsky, Julian Koenig, Anastacia Y. Kudinova, Kevin Kuehn, Peter Kuppens, Anita N.D. Kwashie, Michael J. Kyron, Shon W. Lewis, Paul S. Links, Richard T. Liu, Jorge López-Castroman, Kevin M. Malone, Juan Martínez-Miranda, Ivan W. Miller, Regina Miranda, Jennifer J. Muehlenkamp, Maria A. Oquendo, Andrew C. Page, David Pagliaccio, Jasper E. Palmier-Claus, Evyn M. Peters, Wilfred R. Pigeon, Mina M. Rizk, Heather T. Schatten, Lori N. Scott, Edward A. Selby, Jennifer S. Silk, Ellen Sleuwaegen, Joshua M. Smyth, Lena Spangenberg, Barbara H. Stanley, Jana-sophie Stenzel, Stephanie D. Stepp, Lana Stermac, Gregory L. Stuart, Brianna J. Turner, Elizabeth Valencia, Kristof Vansteelandt, Daniel Vega, Sarah E. Victor, Vera Vine, Hilary Weingarden and Caitlin Wolford-Clevenger.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cpr.2021.102098>.

### References

- Bai, S., Babeva, K. N., Kim, M. I., & Asarnow, J. R. (2020). Future directions for optimizing clinical science & safety: Ecological momentary assessments in suicide/self-harm research. *Journal of Clinical Child & Adolescent Psychology*, 1–13. <https://doi.org/10.1080/15374416.2020.1815208>
- Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change*. Allyn and Bacon.
- Bentley, K. H., Kleiman, E. M., Elliott, G., Huffman, J. C., & Nock, M. K. (2019). Real-time monitoring technology in single-case experimental design research: Opportunities and challenges. *Behaviour Research and Therapy*, 117, 87–96. <https://doi.org/10.1016/j.brat.2018.11.017>
- Berman, A. L., & Carter, G. (2020). Technological advances and the future of suicide prevention: Ethical, legal, and empirical challenges. *Suicide and Life-threatening Behavior*, 50(3), 643–651. <https://doi.org/10.1111/sltb.12610>
- Briere, J., & Gil, E. (1998). Self-mutilation in clinical and general population samples: Prevalence, correlates, and functions. *American Journal of Orthopsychiatry*, 68(4), 609–620. <https://doi.org/10.1037/h0080369>
- Buehler, J. W. (2017). Racial/Ethnic disparities in the use of lethal force by US police, 2010–2014. *American Journal of Public Health*, 107(2), 295–297. <https://doi.org/10.2105/AJPH.2016.303575>
- Centers for Disease Control and Prevention. (2020). Preventing suicide. [https://www.cdc.gov/violenceprevention/pdf/Suicide-factsheet\\_508.pdf](https://www.cdc.gov/violenceprevention/pdf/Suicide-factsheet_508.pdf).
- Coppersmith, D. D. L., Fortgang, R., Kleiman, E., Millner, A., Yeager, A., Mair, P., & Nock, M. (2020). Frequent assessment of suicidal thinking does not increase suicidal thinking: Evidence from a high-resolution real-time monitoring study [preprint]. *PsyArXiv*. <https://doi.org/10.31234/osf.io/6bh82>
- Edwards, F., Lee, H., & Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race-ethnicity, and sex. *Proceedings of the National Academy of Sciences*, 116(34), 16793–16798. <https://doi.org/10.1073/pnas.1821204116>
- Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., Musacchio, K. M., Jaroszewski, A. C., Chang, B. P., & Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin*, 143(2), 187–232. <https://doi.org/10.1037/bul0000084>
- Gee, B. L., Han, J., Benassi, H., & Batterham, P. J. (2020). Suicidal thoughts, suicidal behaviours and self-harm in daily life: A systematic review of ecological momentary assessment studies. *DIGITAL HEALTH*, 6. <https://doi.org/10.1177/2055207620963958>, 2055207620963958.
- Glenn, C. R., & Nock, M. K. (2014). Improving the short-term prediction of suicidal behavior. *American Journal of Preventive Medicine*, 47(3 Suppl 2), S176. <https://doi.org/10.1016/j.amepre.2014.06.004>
- Hepp, J., Carpenter, R. W., Störkel, L. M., Schmitz, S. E., Schmah, C., & Nijfeld, I. (2020). A systematic review of daily life studies on non-suicidal self-injury based on the four-function model. *Clinical Psychology Review*, 82, Article 101888. <https://doi.org/10.1016/j.cpr.2020.101888>
- Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychology*, 34(Suppl), 1220–1228. <https://doi.org/10.1037/hea0000305>
- Kleiman, E. M., Millner, A. J., Kellerman, J. K., Yeager, A. L., Grove, J. G., Bentley, K. H., Nash, C. C., & Nock, M. K. (2021). A qualitative assessment of adolescents' and their parents' preferences on responding to suicide risk in intensive longitudinal monitoring studies. Manuscript in preparation.
- Kleiman, E. M., & Nock, M. K. (2017). Real-time assessment of suicidal thoughts and behaviors. *Current Opinion in Psychology*, 22, 33–37.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7). <https://doi.org/10.1371/journal.pmed.1000097>
- Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-Time Adaptive Interventions (JITAs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6), 446–462. <https://doi.org/10.1007/s12160-016-9830-8>
- National Institute of Mental Health. Conducting research with participants at elevated risk for suicide: Considerations for researchers. Retrieved December 14, 2020, from <https://www.nimh.nih.gov/funding/clinical-research/conducting-research-with-participants-at-elevated-risk-for-suicide-considerations-for-researchers.shtml#regulatory>
- Nock, M., Joiner, T., Gordon, K., Lloydrichardson, E., & Prinstein, M. (2006). Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry Research*, 144(1), 65–72. <https://doi.org/10.1016/j.psychres.2006.05.010>
- Nock, M. K., Kleiman, E. M., Abraham, M., Bentley, K. H., Brent, D. A., Buonopane, R. J., ... Pearson, J. L. (2020). Consensus statement on ethical & safety practices for conducting digital monitoring studies with people at risk of suicide and related behaviors. *Psychiatric Research and Clinical Practice*. <https://doi.org/10.1176/appi.prcp.20200029>. appi.prcp.20200029.
- Pew Research Center. (2019). Mobile Fact Sheet. <https://www.pewresearch.org/internet/fact-sheet/mobile/>.
- Rabasco, A., & Sheehan, K. (2021). The use of intensive longitudinal methods in research on suicidal thoughts and behaviors: A systematic review. *Archives of Suicide Research*, 1–15. <https://doi.org/10.1080/13811118.2021.1903635>
- Reynolds, B. M., Robles, T. F., & Repetti, R. L. (2016). Measurement reactivity and fatigue effects in daily diary research with families. *Developmental Psychology*, 52(3), 442–456. <https://doi.org/10.1037/dev0000081>

- Ribeiro, J. D., Franklin, J. C., Fox, K. R., Bentley, K. H., Kleiman, E. M., Chang, B. P., & Nock, M. K. (2016). Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: A meta-analysis of longitudinal studies. *Psychological Medicine*, *46*(2), 225. <https://doi.org/10.1017/S0033291715001804>
- Rodríguez-Blanco, L., Carballo, J. J., & Baca-García, E. (2018). Use of ecological momentary assessment (EMA) in non-suicidal self-injury (NSSI): A systematic review. *Psychiatry Research*, *263*, 212–219. <https://doi.org/10.1016/j.psychres.2018.02.051>
- Schatten, H. T., Gaudiano, B. A., Primack, J. M., Arias, S. A., Arney, M. F., Miller, I. W., Epstein-Lubow, G., & Weinstock, L. M. (2020). Monitoring, assessing, and responding to suicide risk in clinical research. *Journal of Abnormal Psychology*, *129*(1), 64–69. <https://doi.org/10.1037/abn0000489>
- Sedano-Capdevila, A., Porras-Segovia, A., Bello, H. J., Baca-García, E., & Barrigon, M. L. (2021). Use of ecological momentary assessment to study suicidal thoughts and behavior: A systematic review. *Current Psychiatry Reports*, *23*(7), 41. <https://doi.org/10.1007/s11920-021-01255-7>
- Shepard, D. S., Gurewich, D., Lwin, A. K., Reed, G. A., & Silverman, M. M. (2016). Suicide and suicidal attempts in the United States: Costs and P. *Suicide and Life-threatening Behavior*, *46*(3), 352–362. <https://doi.org/10.1111/sltb.12225>
- Swannell, S. V., Martin, G. E., Page, A., Hasking, P., & St John, N. J. (2014). Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression. *Suicide and Life-threatening Behavior*, *44*(3), 273–303. <https://doi.org/10.1111/sltb.12070>
- Turner, B. J., Austin, S. B., & Chapman, A. L. (2014). Treating nonsuicidal self-injury: A systematic review of psychological and pharmacological interventions. *The Canadian Journal of Psychiatry*, *59*(11), 576–585. <https://doi.org/10.1177/070674371405901103>
- van Ballegooijen, W., Ruwaard, J., Karyotaki, E., Ebert, D. D., Smit, J. H., & Riper, H. (2016). Reactivity to smartphone-based ecological momentary assessment of depressive symptoms (MoodMonitor): Protocol of a randomised controlled trial. *BMC Psychiatry*, *16*(1), 359. <https://doi.org/10.1186/s12888-016-1065-5>