

# Bridging the gap from medical to psychological safety assessment: consensus study in a digital mental health context

Rayan Taher, Palak Bhanushali, Stephanie Allan, Mario Alvarez-Jimenez, Heather Bolton, Laura Dennison, Brian E. Wallace, Heather D. Hadjistavropoulos, Charlotte L. Hall, Amy Hardy, Alasdair L. Henry, Sam Lane, Tess Maguire, Adam Moreton, Talar R. Moukhtarian, Elvira Perez Vallejos, Sukhi Shergill, Daniel Stahl, Graham R. Thew, Ladislav Timulak, David van den Berg, Noemi Viganò, Ben Wensley Stock, Katherine S. Young and Jenny Yiend

## Background

Digital Mental Health Interventions (DMHIs) that meet the definition of a medical device are regulated by the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK. The MHRA uses procedures that were originally developed for pharmaceuticals to assess the safety of DMHIs. There is recognition that this may not be ideal, as is evident by an ongoing consultation for reform led by the MHRA and the National Institute for Health and Care Excellence.

## Aims

The aim of this study was to generate an experts' consensus on how the medical regulatory method used for assessing safety could best be adapted for DMHIs.

## Method

An online Delphi study containing three rounds was conducted with an international panel of 20 experts with experience/knowledge in the field of UK digital mental health.

## Results

Sixty-four items were generated, of which 41 achieved consensus (64%). Consensus emerged around ten recommendations, falling into five main themes: Enhancing the quality of adverse

events data in DMHIs; Re-defining serious adverse events for DMHIs; Reassessing short-term symptom deterioration in psychological interventions as a therapeutic risk; Maximising the benefit of the Yellow Card Scheme; and Developing a harmonised approach for assessing the safety of psychological interventions in general.

## Conclusion

The implementation of the recommendations provided by this consensus could improve the assessment of safety of DMHIs, making them more effective in detecting and mitigating risk.

## Keywords

Digital mental health interventions; mental health; safety; MHRA; consensus.

## Copyright and usage

© The Author(s), 2024. Published by Cambridge University Press on behalf of Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Digital Mental Health Interventions (DMHIs) are interventions that use technologies such as apps, computers or virtual reality to support people living with mental health conditions.<sup>1,2</sup> Such interventions have the potential to reach a large number of individuals, significantly widen access to care and effectively scale evidence-based support.<sup>1</sup> There is increasing evidence for the effectiveness of DMHIs in improving mental health conditions; however, evaluating their safety remains a key challenge for professionals in the field.<sup>1,3,4</sup>

DMHIs that meet the definition of a 'medical device' are regulated by various medical regulatory bodies across the world.<sup>5,6</sup> A medical device, as defined by the EU Directive 93/42/EEC.11, is 'any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, including the software intended by its manufacturer to be used specifically for diagnostic and/or therapeutic purposes and necessary for its proper application, intended by the manufacturer to be used for human beings for the purpose of diagnosis, prevention, monitoring, treatment or alleviation of disease'.<sup>7</sup> In the UK, the Medicines and Healthcare products Regulatory Agency (MHRA) is responsible for regulating both pharmaceuticals and medical devices, and thus all DMHIs are classified as such.<sup>6</sup> Therefore the UK's regulatory framework currently applies only to DMHIs intended to treat or alleviate mental ill health, for example those designed for specific, clinically recognised conditions like depression, paranoia or

insomnia and specifically intended for use only in those predefined populations. In contrast there is a wide range of similar products that are intended to improve mental well-being but which fall outside the medical regulatory framework and therefore the remit of this paper.<sup>6,8</sup> This includes anything designed for general public use, for example wellness, exercise, healthy eating, meditation/mindfulness or motivational apps.

From a safety perspective, a medical device's life cycle can be divided into two phases: the pre- and post-market phase. The pre-market phase is when the product's effectiveness and safety is still being tested. The post-market phase is when the product has been approved by the appropriate regulators to be placed on the market and used publicly.<sup>6</sup>

## Assessing the safety of DMHIs in the pre-market phase

In the pre-market phase, the safety of DMHIs is assessed mainly in clinical trials by collecting and analysing Adverse Events and Serious Adverse Events (SAEs) data.<sup>9</sup> According to the Good Clinical Practice guidelines of the International Council for Harmonization, an adverse event is an 'untoward medical occurrence', and an SAE is an adverse event that results in hospitalisation, death, disability or a birth defect, is a medically important event or would have resulted in any of these outcomes if preventative action had not been taken.<sup>1</sup>

Negative effects or adverse events relevant to mental health interventions include deterioration of symptoms (the 'deterioration effect'), novel symptoms and/or non-response.<sup>10</sup> The deterioration effect occurs when people's symptoms worsen during therapy.<sup>11</sup> Approximately 3–10% of psychotherapy patients experience deterioration.<sup>11</sup> There is division in the literature on whether symptom deterioration is an integral and necessary part of therapy or whether it is an unwanted side-effect.<sup>12</sup> Novel symptoms are when people experience new mental health symptoms during treatment, ones that they did not present with prior to the onset of the treatment.<sup>13</sup> Non-response is when people's target symptoms do not improve. This is seen as a negative effect because the treatment might have prevented the person from receiving an effective therapy.<sup>13</sup> Furthermore, the digital aspect of DMHIs introduces new risks to mental health therapies such as technical issues and privacy concerns.<sup>1,14</sup>

Despite efforts to adequately assess the safety of mental health interventions, numerous research studies have found that the collection, monitoring and analysis of adverse events and SAEs data in face-to-face/digital psychological therapy trials are still in a premature phase.<sup>4,12–15</sup> A recent review on the identification and categorisation of adverse events in DMHI trials found that only six out of 23 of the included trials (26%) reported adverse events data within their primary results.<sup>1</sup> The lack of consensus on the definition of what constitutes an adverse event and SAE in a trial for a mental health intervention and what data needs to be collected, and how, have been cited as the main reasons for the current inferior standard of adverse events monitoring and reporting in DMHIs.<sup>1,4,16</sup> Furthermore, reliance on guidelines that were originally developed for assessing the safety of pharmacotherapies can be a further hindrance.<sup>1</sup>

### Assessing the safety of DMHIs in the post-market phase

In the post-market phase, the MHRA uses Post Market Surveillance (PMS) to monitor the safety of drugs and medical devices that reach the market.<sup>17</sup> As part of PMS, the manufacturer is responsible for continuously monitoring the safety of their product(s), implementing strategies to minimise risk(s) and reporting their findings to the regulatory body.<sup>7</sup> Additionally, the MHRA uses a spontaneous recording system called the 'Yellow Card Scheme' which is a safety monitoring system.<sup>17</sup> Its purpose is to identify new side-effects that might not have been previously known and to monitor the safety of products on the market. It allows health professionals, as well as patients and caregivers, to report any suspected adverse reactions directly to the MHRA.<sup>17</sup> The Yellow Card Scheme has shown effectiveness in helping the MHRA assess the safety of drugs and medical devices.<sup>17</sup> A review of Yellow Card data found that, compared with health professionals, patients are more likely to report adverse reactions that fall under 'psychiatric disorders' – a categorisation that groups all psychiatric, psychological and mental health reactions under the Yellow Card Scheme.<sup>18</sup> It is speculated that this might be because health professionals may give less importance to such reactions (mental health reactions), or they do not acknowledge such reactions as 'adverse'.<sup>18</sup> This further highlights the importance of adapting and promoting the use of the Yellow Card Scheme to assess the safety of DMHIs.

The regulation of DMHIs using medical regulatory methods has some benefits, such as integrating mental health interventions into an established and robust regulatory framework used across the healthcare sector.<sup>5</sup> However, using a regulatory process for DMHIs that was originally developed for other purposes could have several undesirable consequences, such as missing important harms or overemphasising less significant harms.<sup>4</sup> There is

already an awareness of the need for better regulatory methods for DMHIs; in the UK, this is evident from the £1.8 million received by the MHRA and National Institute for Health and Care Excellence to specifically explore the regulation of DMHIs.<sup>19</sup> To contribute towards this objective, we employed the Delphi technique to reach an experts' consensus on 'What are the adaptations needed for the existing medical based model that is used to assess safety of DMHIs?'

## Method

### Study design

An online Delphi technique was used to reach an experts' consensus on how the medical methods used for assessing safety can be adapted to better meet the objectives of DMHIs. The Delphi technique is a tool used to reach a consensus, through controlled feedback on topics with limited evidence or where a consensus does not exist.<sup>20,21</sup> The main components of a Delphi technique are anonymity, iteration, controlled feedback and statistical stability.<sup>13,21,22</sup>

### Recruitment and participants

Purposive sampling was used to invite participants to participate in this Delphi study. Individuals were eligible to participate if they were professionals (researchers, clinicians or regulatory consultants) with experience/knowledge in the field of digital mental health in the UK. There were no exclusion criteria. Experts who met the eligibility criteria were contacted via email or LinkedIn. A snowball technique was also used to recruit participants. Thirty-six professionals were invited to join the panel, of whom 25 consented (69%) and 20 participated (55%). Thus, the sample size was in the average range for Delphi studies.<sup>20</sup> Of the 20 participants, 11 (55%) were females. The majority of the participants (17/20, 85%) were based in the UK, with three living in Canada, Australia and the Netherlands. Participants had an average of 19 years' (range 5–37, s.d. = 7.45) general working experience and 8 years (range 2–15, s.d. = 4.51) in the field of digital mental health. See Table 1 for further details on the participants' occupations. As common in expert Delphi studies, participating panel members were offered authorship on any resulting publications.<sup>23</sup>

### Data collection and analysis

Three rounds were conducted to reach a consensus in this study. Participants were given 2 weeks to complete each round.

Round 1: item generation

Participants were sent a link to an 11-question qualitative questionnaire to collect their opinion on how the safety of DMHIs is currently assessed (see Supplementary Appendix A available at <https://doi.org/10.1192/bjo.2024.713> for the list of questions).

**Table 1** Experts' occupations and employment sector

	Number	Percentage
Experts' occupation		
Academic	7	35
Clinical academic	7	35
Clinician	4	20
Regulatory consultant	1	5
Researcher in the industry (doctoral level)	1	5
Experts' sectors of employment		
Higher education	13	65
Industry (private sector)	5	25
Regulatory board	2	10

Open-ended qualitative questionnaires are usually used in Delphi studies to generate items using the panel's opinions.<sup>20</sup> Twenty participants completed round 1. The qualitative data collected were analysed using open coding to generate items. Open coding is 'the process of identifying, naming, and categorising concepts in the data'.<sup>24</sup> The authors R.T. and P.B. analysed the data separately and generated items based on the data. Afterwards, they compared and discussed their individual items and agreed on a final list of 64 items. This process was overseen by a third author J.Y.

Round 2: collecting controlled feedback

A quantitative questionnaire comprising the 64 final items was shared with all 20 participants who participated in round 1. Participants were asked to rate their agreement with each item on a 5-point Likert scale (strongly disagree, disagree, neutral, agree or strongly agree). Twenty participants completed round 2 (100%).

Round 3: reaching a quantitatively predefined consensus

Twenty individual questionnaires were developed. Each questionnaire contained the same items and rating options from round 2, the group's answers, plus the individual's rating. Considering the group's ratings, participants were given an opportunity to reappraise their previous ratings (see Fig. 1). Eighteen participants (90%) completed round 3. The percentage of people who agreed (agree/strongly agree) and disagreed (disagree/strongly disagree) with each item was calculated. Consensus was predefined as any item that reached 80% agreement (the most common method used to define consensus.<sup>25,26</sup>

### Ethical clearance

This study was registered as a minimal risk study at King's College London under the reference number MRSP-22/23-34777. Under the above approval, consent was implied by participants' email agreement to take part and completion of the various requested survey activities (Round 1–3). Furthermore, at the end of each online questionnaire, the following message was presented to participants: 'Please be aware that submitting this form implies consent to participate in this research'.

## Results

Forty-one out of 64 items reached consensus (64%). The items that reached consensus in this study can be grouped under five main themes. Table 2 shows percent agreement for the 41 items reaching consensus, organised by theme. Supplementary Appendix B shows all 64 items with corresponding agreement and disagreement rates.

### Enhancing the quality of adverse events data in DMHIS

Under this theme, experts echoed researchers' findings on the poor quality of adverse events data in DMHI studies.<sup>1</sup> A total of 94% of experts agreed that 'The digital mental health field lacks clear guidance on what adverse events data should be collected, when and how this should happen, and how these should be categorised, analysed and reported'. All experts agreed that 'Generating publication guidelines for reporting adverse events data that are required by journals can help ensure that adverse events data are collected and reported consistently'. Furthermore, under this theme experts concluded that 'Determining the relatedness of the adverse event to the therapy is paramount for assessing safety' (94%) and that 'When assessing adverse events in digital mental health interventions, it is very important to differentiate between self-guided and clinician-guided therapies' (89%).

### Redefining SAEs for DMHIs

One expert said in round 1: 'The current definition of serious adverse events (hospitalisation, death, congenital abnormalities ... ) is not fit for purpose for mental health interventions compared to pharmaceutical, surgical etc.'. When this item was later shared with the panel of experts in round 3, 94% concurred with it. In all, 83% of experts agreed that 'Congenital abnormalities as a serious adverse event of digital mental health interventions is not plausible'. Experts also expressed that the traditional definition of an SAE is not comprehensive of what should be considered an SAE for a DMHI. All experts agreed that 'In the context of mental health therapies, it is important for "serious" to be inclusive of serious effects on mood, behaviour, general well-being and functioning, not just physical harms'. In addition to loss of life, hospitalisation and disability, suicide attempts (94%), increase in harm towards others (100%), increase in intentional self-harm (100%)

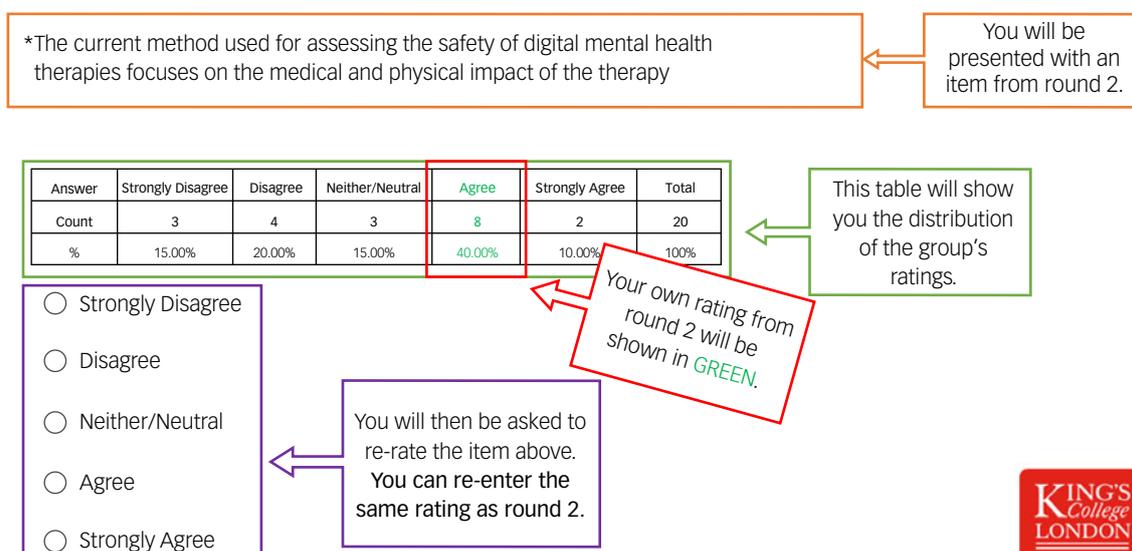


Fig. 1 A screenshot of the guide used to assist participants in round 3.

**Table 2** Items from round 1 reaching consensus ( $\geq 80$  agreement) organised by theme. Bold text denotes items scoring  $>90\%$  agreement

Theme	Generated item (percentage of agreement)
Enhancing the quality of adverse events data in Digital Mental Health Interventions (DMHIs)	<b>Generating publication guidelines for reporting adverse events data that are required by journals can help ensure that adverse events data are collected and reported consistently.</b> (100%)
	<b>Determining the relatedness of the adverse event to the therapy is paramount for assessing safety.</b> (94%)
Redefining serious adverse events for DMHIs	<b>The digital mental health field lacks clear guidance on what adverse events data should be collected, when and how this should happen, and how these should be categorised, analysed and reported.</b> (94%)
	When assessing adverse events in DMHIs, it is very important to differentiate between self-guided and clinician-guided therapies. (89%)
	<b>In the context of mental health therapies, it is important for 'serious' to be inclusive of serious effects on mood, behaviour, general well-being and functioning, not just physical harms.</b> (100%)
	<b>Increase in harm towards others.</b> (100%)
	<b>Increase in intentional self-harm.</b> (100%)
	<b>Seizures induced because of on-screen flashes.</b> (100%)
	<b>A suicide attempt.</b> (94%)
	<b>The current definition of serious adverse events (hospitalisation, death, congenital abnormalities ...) is not fit for purpose for mental health interventions compared with pharmaceutical, surgical etc.</b> (94%)
	Any harm that is sustained for longer than 6 months. (89%)
	Congenital abnormalities as a serious adverse event of digital mental health intervention is not plausible. (83%)
Reassessing short-term deterioration in face-to-face and DMHIs as a therapeutic risk	<b>Discuss a clear rationale for each therapy activity/exercise.</b> (100%)
	<b>People need to be informed about short-term symptom deterioration/distress (through the process of informed consent). In a digital context, this can be replicated through text, audio or video content.</b> (94%)
	<b>It is important for therapists to normalise symptom deterioration. In a digital context, this can happen as part of a discussion with a trained mental health professional or via the content of the digital therapy.</b> (94%)
	<b>Develop and maintain a strong therapeutic alliance.</b> (94%)
	<b>Help the individual think about sources of support that they can use between sessions if needed. In a digital context, provide people with alternative sources of support should they need it.</b> (94%)
	<b>Provide individuals with information on how to manage symptom deterioration.</b> (94%)
	The short-term distress generated by focusing on the emotional experience (perhaps discussing painful or traumatic experiences or going into new or anxiety-provoking situations) is a part of the therapy process and should not be considered an adverse event or safety concern. (89%)
	Long-term follow-ups are needed to differentiate between short-term and long-term deterioration. (89%)
	Short-term deterioration is a momentary side-effect of therapy and one that should be monitored, so that negative outcomes/harm do not result from the treatment. (83%)
	Monitoring of symptoms through regular symptom checks could also mitigate the risks of deterioration. (83%)
	DMHIs should have pre-set thresholds of deterioration that trigger an escalation such as a face-to-face intervention to mitigate risk. (83%)
	<b>Intentional self-harm.</b> (100%)
	<b>Threat to others.</b> (100%)
	<b>There is not enough awareness among individuals that the Yellow Card System is used for reporting adverse reactions in digital mental health therapies.</b> (100%)
	<b>Grouping mental health adverse events under 'psychiatric disorders' may discourage reporting milder mental health adverse events. Thus, using a broader category such as 'mental health adverse events' is more appropriate.</b> (94%)
<b>Severity/impact on the individual.</b> (94%)	
<b>A key problem seems to be that it relies on individuals to report risks, which they may not be willing or able to do. To mitigate this, Patient Reported Outcomes could be automatically collected within the digital therapy itself.</b> (94%)	
<b>Suicidal ideation/behaviour.</b> (94%)	
Grouping all mental health adverse reactions under 'psychiatric disorders' risks losing useful and important detail on the different mental health adverse events experienced by people which could inform risk mitigation procedures. (89%)	
Whether the event led to the person terminating therapy. (89%)	
A description of the adverse event (free text). (89%)	
Duration of adverse event. (89%)	
New mental health symptoms. (89%)	
Individuals are not aware that the Yellow Card Scheme is used to report the adverse reactions of digital mental health therapies. (83%)	
When categorising adverse events, a broad mental health category is needed with subtypes that relate to the different adverse events that could occur under that category. (83%)	

(Continued)

Table 2 (Continued)

Theme	Generated item (percentage of agreement)
Developing a harmonised approach for assessing the safety of face-to-face and DMHIs	<p><b>Having a harmonised approach for safety reporting across face-to-face and digital interventions will increase our knowledge of the risks of mental health therapies and contribute to stronger safety standards and safer interventions.</b> (94%)</p> <p>Using a harmonised approach to assess safety in face-to-face and digital mental health therapies would allow us to more easily compare the safety of face-to-face and digital therapies. (89%)</p> <p>If there was one structured way to report adverse events in face-to-face and digital mental health therapies, clinicians and mental health professionals would be more likely to report adverse events. (83%)</p> <p>Safety issues that have to do with clinical support need to be harmonised between face-to-face and DMHIs, because of their commonalities. In a digital context, safety issues that arise because of using a digital tool would be added to that. (83%)</p>

and seizures that are induced because of on-screen flashes (100%) should be considered SAEs even if they do not lead to hospitalisation. Experts added that in the case of a DMHI, any harm that is sustained for longer than 6 months should be considered an SAE (89%).

### Reassessing short-term deterioration in face-to-face and DMHIs as a therapeutic risk

Experts reached a consensus that ‘the short-term distress generated by focusing on the emotional experience (perhaps discussing painful or traumatic experiences or going into new or anxiety-provoking situations) is a part of the therapy process and should not be considered an adverse event or safety concern’ (89%). However, they also acknowledged that such a deterioration could lead to negative outcome/harm; thus, such deterioration needs to be monitored and not ignored (83%). For that, experts encouraged long-term follow-ups to differentiate between short-term and long-term deterioration (89%). They also suggested several ways to support individuals experiencing deterioration; eight of these reached consensus (see Supplementary Appendix C).

### Maximising the benefit of the Yellow Card Scheme in DMHIs

All experts agreed that ‘There is not enough awareness among users that the Yellow Card Scheme is used for reporting adverse reactions in digital mental health therapies’. Furthermore, 83% of experts agreed with the item ‘I was not aware that the Yellow Card Scheme is used to report the adverse reactions of digital mental health therapies’. Experts acknowledged that one drawback of the Yellow Card Scheme for collecting adverse reactions is that it relies on individuals to report risks, which they may not be willing or able to do; for that, they suggested that Patient Reported Outcomes could be automatically collected within the digital therapy itself to help identify risks (94%).

Acknowledging the benefit of the Yellow Card Scheme, experts determined a few improvements that would make it more useful for DMHIs. First, experts noted that grouping all mental health adverse reactions under ‘psychiatric disorders’ risks losing useful and important detail on the different mental health adverse reactions experienced by individuals and may discourage reporting milder mental health adverse reactions (89% and 94% respectively). Instead, they suggested having a broader and more inclusive category such as ‘mental health adverse reactions’, with subtypes that relate to the different adverse reactions that could occur under that category (94% and 83% respectively). Experts settled on four proposed subtypes: intentional self-harm (100%), threat to others (100%), suicidal ideation/behaviours (94%) and new mental health symptoms (89%). Additionally, experts agreed that having

the following information about every mental health adverse reaction reported on the Yellow Card website under the DMHI’s profile would be helpful: severity/impact on the individual (94%), duration of adverse reaction (89%), whether the event led to the individual terminating therapy (89%) and a description of the adverse reaction (free text) (89%).

### Developing a harmonised approach for assessing the safety of psychological interventions generally

Experts agreed that there should be a single harmonised approach for assessing and reporting the safety of face-to-face and DMHIs. Experts assented that such an approach would allow professionals to compare the safety of face-to-face and DMHIs (89%), and thus better understand the risks of mental health therapies, and contribute to stronger safety standards and safer interventions (94%). Moreover, experts speculated that ‘If there was one structured way to report adverse reactions in face-to-face and digital mental health therapies, clinicians and mental health professionals will be more likely to report them’ (83%). At the very least, the safety issues related to clinical support that are common to face-to-face and digital therapy such as the deterioration effect need to be harmonised, and the safety issues that arise from the use of a digital tool such as technical issues could be added to that (83%).

### Items with high disagreement rates

We also explored the items that had high rates of disagreement. See Supplementary Appendix B for the percentage of disagreement (i.e. rating disagree/strongly disagree) for all items. No items reached consensus disagreement (80% or higher). However, two items equalled or exceeded 50%. A total of 61% of experts disagreed that ‘triggering trauma-related flashbacks’ should be considered an SAE. Half of the experts disagreed that ‘triggering a panic attack’ should be considered an SAE. Although these items arose directly at the suggestion of some experts, more than half the group disagreed. These disagreement rates illustrate the difficulty defining what constitutes an SAE for DMHIs and reveal that some areas divide opinion.

## Discussion

This study successfully reached an experts’ consensus on ten recommendations for how to adapt the medical method used for assessing the safety of DMHIs. Our panel’s consensus suggestions have the potential to improve the process of safety assessment, enhancing the field’s approach to the safety of DMHIs. Table 3 provides a summary of the recommendations.

**Table 3** List of recommendations on how to adapt the medical regulatory safety assessment model to meet the needs of the digital mental health field

Recommendation 1	Publish guidelines for professionals on how to assess and report adverse events data in digital mental health research.
Recommendation 2	Standardise the optimum way to assess the relatedness of an adverse event to the intervention, procedure or device, by using a variety of independent evidence sources.
Recommendation 3	When evaluating safety, differentiate between self-guided and clinician-guided interventions.
Recommendation 4	Broaden the scope of potential serious adverse events in digital studies to consider impact on mental health.
Recommendation 5	Monitor psychological deterioration to prevent potential harm, using post-treatment follow-ups.
Recommendation 6	Utilise the suggestions outlined (Supplementary Appendix C) to support people experiencing deterioration in any treatment context.
Recommendation 7	Improve the visibility of the Yellow Card Scheme as the primary means of reporting adverse reactions in the digital mental health field.
Recommendation 8	Change the Yellow Card 'psychiatric conditions' category label to 'mental health adverse reactions'.
Recommendation 9 (a & b)	Expand the Yellow Card category that captures mental health data to incorporate informative subcategories and additional contextual information.
Recommendation 10	Develop an overarching, harmonised framework for assessing the safety of any psychological intervention (digital or otherwise).

### Improving adverse events data in DMHI trials

The Delphi study showed a consensus on the lack of clear guidelines on how to collect, analyse and report adverse events data in DMHIs. Using guidelines that were originally developed for pharmacological trials is adversely affecting the quality of adverse events data in DMHIs.<sup>1,16</sup> Consequently, experts unanimously agreed that generating publication guidelines to support professionals on what adverse events data need to be collected, how often and how they need to be analysed, will help standardise adverse events reporting in the field and allow for more comprehensive and systematic analysis of adverse events data (*recommendation 1*).<sup>13</sup>

To effectively assess the safety of DMHIs, determining the relatedness of an adverse event to the treatment is key (*recommendation 2*), which can be less straightforward in a DMHI compared with a pharmacological intervention.<sup>10</sup> Additional data collection to support this judgement could include onset, previous similar instances, and participant's and clinician's opinions on relatedness. This could improve the robustness of relatedness decision-making.<sup>1</sup>

When assessing or comparing the safety of DMHIs, it is important to differentiate between self-guided and clinician-guided DMHIs (*recommendation 3*). The concern is that self-guided interventions can pose more risk in the absence of a professional guiding the individual, explaining content to them, assessing risk and providing emotional support or encouragement to people when they need it.<sup>1</sup> However, a review of the state of adverse events data in DMHIs found that studies of clinician-guided interventions were more likely to address adverse events in their publication compared with self-guided interventions.<sup>1</sup> Adverse events in self-guided interventions are more likely to be missed because of the lack of human contact, and thus might be underreported. Thus, it is crucial for self-guided interventions to integrate within the product a way of collecting and monitoring adverse event and safety data, such as collecting pre/post data to detect deterioration or using automated triggers to alert clinicians to a worsening condition.<sup>26</sup> In such cases, individuals would need to be informed and provide consent for such data to be collected and shared with the clinical team.

### Broadening the definition of SAE for DMHIs

Experts in this study agreed that the pharmacological frame of reference for SAEs is not 'fit for purpose' for DMHIs. They argued that the definition of an SAE in the digital mental health field is too narrow and needs to be inclusive of any potentially life-threatening effects on mood, behaviour, general well-being and functioning. They concurred that adverse events such as increase in harm towards others, increase in intentional self-harm and any harm that is sustained for longer than 6 months should be considered 'serious', regardless of whether it led to hospitalisation or not. An MHRA-led consultation

on the proposed changes to the regulatory framework for medical devices in 2021 (891 respondents) found that professionals requested mental health impacts to be included under 'serious' adverse events.<sup>27</sup> The UK government's response was that at that time 'appropriate mechanisms were not in place to sufficiently regulate the inclusion of these impacts'.<sup>27</sup> We therefore conclude that further consultation is needed to identify and operationalise how the existing SAE definition might best be broadened to capture serious impacts on mental health (*recommendation 4*).

### Short-term deterioration is not a safety concern in mental health therapy

Experts in this study reached a consensus that short-term deterioration of psychological symptoms occurring in the course of a psychological treatment can be an integral part of therapeutic change and is therefore not inevitably a safety concern. Nevertheless, experts stressed that these symptoms could amount to an adverse reaction, for example if the person drops out of therapy as a result. Thus, professionals need to monitor such deterioration and support individuals through it (*recommendation 5*). Similar to other negative effects, individuals need to be informed of the possibility of experiencing deterioration during therapy, and such deterioration needs to be normalised as part of the process.<sup>28</sup> Additionally, guidelines for professionals and patients are needed on the management of short-term deterioration.<sup>29</sup> Our experts provided a list of suggestions (see Supplementary Appendix C) for supporting people experiencing therapy-related deterioration.

### Adapting the Yellow Card Scheme for DMHIs

The medical field has already established that individuals are more likely than healthcare professionals to report adverse reactions that fall under the category 'psychiatric conditions'.<sup>18</sup> Experts in our study reached consensus on several ways to adapt the Yellow Card data collection and reporting process for DMHIs that are classified as medical devices, to maximise their benefit. First, there is a need for a campaign to raise awareness around the use of the Yellow Card Scheme for reporting the adverse reactions to DMHIs among both patients and professionals (*recommendation 7*). One suggestion is to make it a requirement for all DMHIs (including pre-market) to inform and direct people to the Yellow Card reporting website. Second, there was a consensus to use more appropriate, less stigmatising language by replacing 'psychiatric disorders' with 'mental health adverse reactions' (*recommendation 8*). Third, given that these are very general umbrella terms, it would be helpful to include further subcategories. Suggestions from our experts included 'intentional self-harm', 'threat to others', 'suicidal ideation/behaviours' and 'new mental health symptoms'

(*recommendation 9a*). Fourth, it would be helpful to record the following contextual information for every reported adverse reaction: severity/impact on the individual, duration of adverse event, whether therapy was terminated as a consequence and a free text description of the adverse reaction (*recommendation 9b*).

### Using a harmonised approach to assess the safety of psychological therapies generally

Perhaps one of the reasons why it has been difficult to adapt the medical model for DMHIs is that this model was not developed to assess the safety of face-to-face psychological therapies. The current quality of adverse event reporting in trials of face-to-face interventions is poor and is not subject to the same regulation as medical devices.<sup>12</sup> Face-to-face interventions would also benefit from a standardised safety assessment method. Currently there is insufficient information about how the negative effects or adverse events data of face-to-face and DMHIs compare.<sup>13</sup> An ideal solution would be to develop one harmonised approach to assess the safety of face-to-face and DMHIs (*recommendation 10*). One harmonised approach could enhance the quality of safety assessments in traditional face-to-face interventions and ensure that safety evaluations of DMHIs are appropriate, proportionate and effective. Whether it is plausible to meet the needs of statutory regulators, clinicians, researchers, manufacturers and other stakeholders within a single overarching framework remains to be seen. Nevertheless, possible benefits our experts highlighted included better quality safety data, a more accessible method for non-medical mental health practitioners and better safety standards for both face-to-face and DMHIs.

### Limitations

There are several limitations of this Delphi study. First, patients' and their carers' perspectives were absent in the results. Second, 65% of experts who participated work in higher education, leading to a likely overrepresentation of views from within this sector. It is important to note that 9 out of the 16 invitees (56.25%) who were invited but did not participate in the study came from non-academic backgrounds, primarily clinicians and co-founders. Third, evidence reliability is not possible in Delphi studies as the results might have differed if a different group of experts had been selected.

### Conclusion

This Delphi study reached expert consensus on ten recommendations that can be used to adapt the medical regulatory safety assessment model to meet the needs of the digital mental health field. These recommendations, if implemented, have the potential to improve the process of safety assessment of psychological therapies in general and DMHIs in particular. In turn this may result in better quality of mental health safety data, a more effective risk mitigation process in DMHIs and more accurate risk-benefit analysis of DMHIs.

**Rayan Taher** , Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Palak Bhanushali**, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Stephanie Allan** , Institute of Health and Wellbeing, University of Glasgow, UK; **Mario Alvarez-Jimenez**, Centre for Youth Mental Health, University of Melbourne, Australia; and Orygen, Parkville, Australia; **Heather Bolton**, Unmind, London, UK; **Laura Dennison**, School of Psychology, University of Southampton, UK; **Brian E. Wallace**, Calmsie, Warsaw, Poland; **Heather D. Hadjistavropoulos**, Department of Psychology, University of Regina, Canada; **Charlotte L. Hall**, NIHR MindTech-MedTech Co-operative, NIHR Nottingham Biomedical Research Centre, School of Medicine, Institute of Mental Health, University of Nottingham, UK; **Amy Hardy**, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Alasdair L. Henry** , Big Health Ltd, London, UK; **Sam Lane**, SilverCloud by

Amwell, Boston, USA; **Tess Maguire**, School of Psychology, University of Southampton, UK; **Adam Moreton**, ORCHA, Daresbury, UK; **Talar R. Moukhtarian**, Mental Health and Wellbeing Unit, Warwick Medical School, University of Warwick, UK; **Elvira Perez Vallejos**, NIHR MindTech-MedTech Co-operative, NIHR Nottingham Biomedical Research Centre, School of Medicine, Institute of Mental Health, University of Nottingham, UK; **Sukhi Shergill**, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; and Kent and Medway Medical School, Canterbury, UK; **Daniel Stahl**, Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Graham R. Thew**, Department of Experimental Psychology, University of Oxford, UK; and Oxford Health NHS Foundation Trust, Oxford, UK; **Ladislav Timulak**, School of Psychology, Trinity College Dublin, Ireland; **David van den Berg**, Department of Clinical Psychology, VU University and Amsterdam Public Health Research, Amsterdam, Netherlands; **Noemi Viganò**, SilverCloud by Amwell, Boston, USA; **Ben Wensley Stock**, University of Oxford Medical Sciences Division, University of Oxford, UK; **Katherine S. Young**, SilverCloud by Amwell, Boston, USA; and Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Jenny Yiend**, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

**Correspondence:** Jenny Yiend. Email: [jenny.yiend@kcl.ac.uk](mailto:jenny.yiend@kcl.ac.uk)

First received 27 Nov 2023, final revision 17 Apr 2024, accepted 22 Apr 2024

### Supplementary material

Supplementary material is available online at <https://doi.org/10.1192/bjo.2024.713>

### Data availability

The data that support the findings of this study are available on request from the corresponding author, J.Y.

### Acknowledgements

This work was supported by the Medical Research Council (MRC) Biomedical Catalyst: Developmental Pathway Funding Scheme, MRC Reference: MR/V027484/1. We would like to express our gratitude to all the experts who participated in this study, helping us contribute to the development of the field. We also would like to express our gratitude to the National Institute for Health and Care Research (NIHR) Biomedical Research Centre hosted at South London and Maudsley NHS Foundation Trust in partnership with King's College London. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, Department of Health and Social Care, the ESRC or King's College London. For the purposes of open access, the corresponding author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising from this submission.

### Author contributions

R.T., D.S., S.S. and J.Y. designed the study. R.T. applied for ethics, collected data for the study, analysed it and wrote the first draft of the study under the supervision of D.S., S.S. and J.Y. P.B. was involved in the data analysis. S.A., M.A.-J., H.B., L.D., B.E.W., H.D.H., C.L.H., A.H., A.L.H., S.L., T.M., A.M., T.R.M., E.P.V., G.R.T., L.T., D.v.d.B., N.V., B.W.S. and K.S.Y. were all participants in the study and were involved in the write-up process.

### Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

### Declaration of interest

None of the authors have any interests to declare, except A.L.H. who reports being employed by Big Health Ltd and is a shareholder in the company.

### References

- Gómez Bergin AD, Valentine AZ, Rennick-Egglestone S, Slade M, Hollis C, Hall CL. Identifying and categorizing adverse events in trials of digital mental health interventions: narrative scoping review of trials in the international standard randomized controlled trial number registry. *JMIR Ment Health* 2023; **10**: e42501.
- Torous J, Firth J, Mueller N, Onnela JP, Baker JT. Methodology and reporting of mobile health and smartphone application studies for schizophrenia. *Harv Rev Psychiatry* 2017; **25**(3): 146–54.
- Martinez-Martin N, Kreitmair K. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. *JMIR Ment Health* 2018; **5**(2): e9423.
- Taher R, Hsu CW, Hampshire C, Fialho C, Heaysman C, Stahl D, et al. The safety of digital mental health interventions: systematic review and recommendations. *JMIR Ment Health* 2023; **10**(1): e47433.

- 5 Medicines and Healthcare Products Regulatory Agency. *Notify the MHRA about a Clinical Investigation for a Medical Device*. UK Government, 2014 (<https://www.gov.uk/guidance/notify-mhra-about-a-clinical-investigation-for-a-medical-device>).
- 6 Carl JR, Jones DJ, Lindhiem OJ, Doss BD, Weingardt KR, Timmons AC, et al. Regulating digital therapeutics for mental health: opportunities, challenges, and the essential role of psychologists. *Br J Clin Psychol* 2022; **61**(S1): 130–5.
- 7 Parvizi N, Woods K. Regulation of medicines and medical devices: contrasts and similarities. *Clin Med* 2014; **14**(1): 6.
- 8 Espie CA, Henry AL. Disseminating cognitive behavioural therapy (CBT) for insomnia at scale: capitalising on the potential of digital CBT to deliver clinical guideline care. *J Sleep Res* 2023; **32**(6): e14025.
- 9 Wykes T, Lipshitz J, Schueller SM. Towards the design of ethical standards related to digital mental health and all its applications. *Curr Treat Opt Psychiatry* 2019; **6**: 232–42.
- 10 Rozental A, Kottorp A, Boettcher J, Andersson G, Carlbring P. Negative effects of psychological treatments: an exploratory factor analysis of the negative effects questionnaire for monitoring and reporting adverse and unwanted events. *PLoS One* 2016; **11**(6): e0157503.
- 11 Berk M, Parker G. The elephant on the couch: side-effects of psychotherapy. *Aust N Z J Psychiatry* 2009; **43**(9): 787–94.
- 12 Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. *Clin Psychol Psychother* 2013; **20**(4): 286–96.
- 13 Rozental A, Andersson G, Boettcher J, Ebert DD, Cuijpers P, Knaevelsrud C, et al. Consensus statement on defining and measuring negative effects of internet interventions. *Internet Interv* 2014; **1**: 12–9.
- 14 Bradstreet S, Allan S, Gumley A. Adverse event monitoring in mHealth for psychosis interventions provides an important opportunity for learning. *J Ment Health* 2019; **28**: 461–6.
- 15 Terry NP, Gunter TD. Regulating mobile mental health apps. *Behav Sci Law* 2018; **36**(2): 136–44.
- 16 Papaioannou D, Cooper C, Mooney C, Glover R, Coates E. Adverse event recording failed to reflect potential harms: a review of trial protocols of behavioural, lifestyle and psychological therapy interventions. *J Clin Epidemiol* 2021; **136**: 64–76.
- 17 Raj N, Fernandes S, Charyulu NR, Dubey A, Ravi GS, Hebbar S. Postmarket surveillance: a review on key aspects and measures on the effective functioning in the context of the United Kingdom and Canada. *Ther Adv Drug Saf* 2019; **10**: 204209861986541.
- 18 Hazell L, Cornelius V, Hannaford P, Shakir S, Avery AJ. How do patients contribute to signal detection? A retrospective analysis of spontaneous reporting of adverse drug reactions in the UK's yellow card scheme. *Drug Saf* 2013; **36**(3): 199–206.
- 19 Medicines and Healthcare Products Regulatory Agency. *Mental Health Funding of £1.8 m Welcomed by MHRA and NICE to Explore Regulation of Digital Mental Health Tools*. UK Government, 2022 (<https://www.gov.uk/government/news/mental-health-funding-of-18m-welcomed-by-mhra-and-nice-to-explore-regulation-of-digital-mental-health-tools>).
- 20 Taylor E. We agree, don't we? The Delphi method for health environments research. *Health Environ Res Design J* 2020; **13**(1): 11–23.
- 21 Santaguida P, Dolovich L, Oliver D, Lamarche L, Gilsing A, Griffith LE, et al. Protocol for a Delphi consensus exercise to identify a core set of criteria for selecting health related outcome measures (HROM) to be used in primary health care. *BMC Fam Pract* 2018; **19**(1): 1–4.
- 22 Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. *World J Methodol* 2021; **11**(4): 116–29.
- 23 Smoktunowicz E, Barak A, Andersson G, Banos RM, Berger T, Botella C, et al. Consensus statement on the problem of terminology in psychological interventions using the internet or digital components. *Internet Interv* 2020; **21**: 100331.
- 24 Cascio MA, Lee E, Vaudrin N, Freedman DA. A team-based approach to open coding: considerations for creating intercoder consensus. *Field Methods* 2019; **31**(2): 116–30.
- 25 Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 2014; **67**(4): 401–9.
- 26 Seidler ZE, Rice SM, Ogrodniczuk JS, Oliffe JL, Shaw JM, Dhillon HM. Men, masculinities, depression: implications for mental health services from a Delphi expert consensus study. *Prof Psychol Res Pr* 2019; **50**(1): 51–61.
- 27 Medicines and Healthcare Products Regulatory Agency. *Government Response to Consultation on the Future Regulation of Medical Devices in the United Kingdom*. MHRA, 2022 ([https://assets.publishing.service.gov.uk/media/62b577f6d3bf7f0b00165a32/Government\\_response\\_to\\_consultation\\_on\\_the\\_future\\_regulation\\_of\\_medical\\_devices\\_in\\_the\\_United\\_Kingdom.pdf](https://assets.publishing.service.gov.uk/media/62b577f6d3bf7f0b00165a32/Government_response_to_consultation_on_the_future_regulation_of_medical_devices_in_the_United_Kingdom.pdf)).
- 28 Linden M, Schermuly-Haupt ML. Definition, assessment and rate of psychotherapy side effects. *World Psychiatry* 2014; **13**: 306–9.
- 29 Curran J, Parry GD, Hardy GE, Darling J, Mason AM, Chambers E. How does therapy harm? A model of adverse process using task analysis in the meta-synthesis of service users' experience. *Front Psychol* 2019; **10**: 347.

