



Discussion Paper

# Reimagining Digital Mental Health in Australia



**Black Dog**  
Institute



**BETTER  
MENTAL  
HEALTH**

October 2023

**Science.  
Compassion.  
Action.**

# Contents

|  |           |
|--|-----------|
| <b>Chapter 1.</b>  | <b>4</b>  |
| Using technology to build resilient communities            |           |
| <hr/>  |           |
| <b>Chapter 2.</b>  | <b>8</b>  |
| Are screens making young people sad and anxious?           |           |
| <hr/>  |           |
| <b>Chapter 3.</b>  | <b>12</b> |
| Blended care: combining human and digital in clinical care |           |
| <hr/>  |           |
| <b>Chapter 4.</b>  | <b>16</b> |
| Rise of AI in mental health – risks and opportunities      |           |
| <hr/>  |           |

## Acknowledgements

Black Dog Institute acknowledges the Traditional Custodians of the lands we are on, the Gadigal people of the Eora Nation, and pays respect to their Elders past, present and emerging. Black Dog Institute recognises and respects their cultural heritage, beliefs and relationship to their ancestral lands, which continue to be important to First Nations peoples living today.

**Authors:** Aliza Werner-Seidler, Jill Newby, Peter Baldwin, Sam Harvey.

**Editors:** Nicole Scott, Lawrence Muskitta, Harry Grant.

**Design:** Riley Pierce

**Reference:** Black Dog Institute (2023). Reimagining Digital Mental Health in Australia. Sydney, AU: Black Dog Institute



# Foreword

As technology progresses, our mental and digital worlds become increasingly intertwined. Our devices extend our nervous systems through fibre-optic and wireless signals, connecting us with people and places we may never physically touch. There is a seeming paradox when we think about the effects of this new form of connection on mental health.

On one hand, digital technologies have given rise to new clinical and community-based approaches to detect, prevent and treat mental health concerns. Awareness of mental health is at an all-time high, largely due to people sharing their experiences and building communities through digital channels. Similarly, adoption of telehealth has skyrocketed in the past few years, as a response to COVID lockdowns, providing consumers with greater access and choice of mental health services, regardless of where they live.

On the other hand, there is a sense that these same technologies, designed to connect us, are at least partly responsible for a growing feeling of loneliness in individuals and the steep rise of polarisation in our social and political spheres. Despite increased awareness of mental health and the development of new digital therapeutics, rates of mental ill health and suicide continue to rise.

This Discussion Paper is written at an inflection point in digital mental health research and practice. The science we are working on today is still in its nascent phases and we are only scratching the surface of what's possible with mind-digital interfaces. New breakthroughs are being made every day that challenge what we know and question current paradigms. What we can provide is an overview of the evidence as it now stands and our ideas on future directions.

In the following chapters we ask:

- Can we use technology to build more resilient communities?
- Are screens making Australian young people sad and anxious?
- How can we better blend digital and human supports in clinical care?
- What are the risks and opportunities of AI in mental health?

As you read these chapters, we challenge you to approach the content with an open mind, hold the tension of paradox and to ask your own questions. What are we missing? What other approaches or experiences should we be exploring? How can we be working better together to find these answers? How can we turn these insights into action and outcomes?

From a policy perspective, it's clear that more needs to be done to incentivise responsible innovation for digital mental health products and to encourage the uptake of evidence-based digital interventions. Our hope is that this paper will spark inspiration and conversation, as together we reimagine and shape digital mental health in Australia.



# 1. | Using technology to build resilient communities

Humans have evolved to be resilient. There are countless examples of people being able to thrive despite demanding circumstances and to be able to 'bounce back' from adversity. Sadly, there is no shortage of adversity in our community at present. The cost-of-living crisis is placing many families under enormous stress, at a time when many communities were still trying to recover from droughts, bushfires, floods and the impact of COVID-19. We also know that further major natural disasters are very likely. We have never needed resilience more, but at the same time, the building blocks of resilience are moving. The consequences of allowing our resilience to slip at this moment in time will be disastrous. Given this rising burden of cumulative trauma in our communities, we must establish ways to understand and then increase resilience. Without doing this the trend of increasing mental ill health will continue and potentially accelerate.

## What is resilience?

There have been many different definitions of resilience proposed over recent decades. Some researchers have described it as the ability to adapt positively to stressful circumstances<sup>1</sup>, while others have defined resilience as being able to remain functionally stable and well despite ongoing stress<sup>2</sup>. In more recent years, one of the more popular definitions of resilience has come from the American Psychological Society (APA), who define resilience as a process of "bouncing back" from difficult experiences and "adapting well in the face of adversity, trauma, tragedy, threats or significant sources of stress"<sup>3</sup>.

Regardless of exactly how it is defined, a growing number of studies have found a positive link between psychological resilience and mental health outcomes. More specifically, higher resilience levels have been associated with

lower levels of anxiety, psychological distress and depression, especially after traumatic experiences<sup>4</sup>. This is why resilience has become a topic of such interest to the mental health field. If there are ways that resilience can be increased, this could be the key to preventing mental health problems.

## Who has resilience and why?

Before considering how resilience may be able to be enhanced, it is important to consider where resilience resides. It is a mistake to think that resilience is an individual trait. Individual factors are important, but increasingly we understand that resilience is best considered as operating within groups or communities. The Harvard Study of Adult Development is the longest ever study of human resilience, having followed people for over 80 years. The number one finding from all this data is that close relationships and social

1 Luthar, 2000

2 Bonanno, 2004

3 Comas-Diaz, 2016

4 Sood et al., 2011; Loprinzi et al., 2011; Steinhardt and Dolbier, 2008; Connor, 2003

connectiveness is crucial for happiness and resilience<sup>5</sup>. Therefore, if we want to have more resilient communities, communities that will be able to bounce back and not break when the next natural disaster hits, we have to find ways to facilitate and promote social connectiveness.

While social connection may be the most important predictor of resilience in a community, it is not the only lever we have at our disposal. There is increasing evidence that individual factors involved in resilience can also be modified. A large systematic review we recently published on this topic showed that a variety of techniques, including mindfulness and cognitive behavioural training, could generate meaningful shifts in resilience<sup>6</sup>.

## Can technology help build resilience communities?

When a natural disaster or other adversity hits a community, a reasonable question governments and policy makers will ask is what can be done to enhance that community's resilience. Part of the answer to this type of question has to be that the groundwork for resilience needs to have been laid prior to adversity. The social connections need to already be in place and young people should have been taught key coping skills as part of their education. While this is obviously true, it is not the whole story.

In the aftermath of the Black Summer Bushfires and during the COVID-19 pandemic, we have been able to demonstrate that technology can be used to rapidly roll out resilience enhancing

interventions at scale.

Perhaps the most obvious example of how technology can be used in a scalable way to deliver evidence-based resilience training. At the height of COVID related lockdowns, many Australians were feeling stressed and overwhelmed. We conducted a study that used videoconferencing to deliver a brief, low intensity training program over six sessions. The results were clear – this type of technology enabled training was feasible and resulted in substantial reductions in distress and prevented mental health symptoms developing over the next six months. Similarly, we are currently testing the effectiveness of resilience training that is delivered via a smartphone app in disaster impacted regional areas.

This type of training is not the only way technology can be used to build and maintain more resilient communities after adversity. Social media and video conferencing can be used to help maintain social connections, even when situations do not allow face to face interactions. Technology can also be used to ensure that people receive interventions as early as possible. Following the Black Summer Bushfires and the 2022 floods, the Black Dog Institute and UNSW established the National Emergency Worker Support Service (NEWSS). Given the regional location of many of the emergency service workers and volunteers involved in responding to these disasters, NEWSS had to use technology to reach those in need. Since its launch three years ago, this technologically enabled service

---

5 Waldinger R & Schulz M. 2023

6 Joyce et al., 2018

has allowed more than 15,000 emergency service workers to have an online Mental Health Checks and hundred to have treatment for conditions such as PTSD delivered via videoconferencing

have been completed. To date over 80% of people receiving PTSD treatment have made a full and lasting recovery.



## Discussion questions

1. Do you think Australian communities are less resilient than they used to be? If so, why has this happened?
2. The concept of resilience has some similarities with the First Nations concept of social and emotional wellbeing. What can we learn from this concept to enhance our understanding of resilience?
3. Are we kidding ourselves to think that technology can improve resilience? If the key to community resilience is social connection, then isn't face to face communication and education the only way to increase resilience?
4. How should we use technology if Australia has a severe bushfire season? What can we be doing now to prepare?
5. How might technology be being used to improve resilience in ten years' time? Are there exciting developments coming soon?

## References

- Black Dog Institute. 2022. "Turning the Tide on Depression." White Paper.
- Bonanno GA. 2004. "Loss, trauma, and human resilience." *American Psychologist* 59: 20-28.
- Comas-Díaz L, Luhtar SS, Maddi SR, O'Neill HK, Saakvithe KW, & Tedeschi RG. 2016. "The Road to Resilience." In: APA (ed). <http://www.apa.org/helpcenter/road-resilience.aspx>.
- Connor KM, & Davidson JRT. 2003. "Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC)." *Depression and Anxiety* 18: 71-82.
- Joyce S, Shand F, Tighe J, et al. 2018. "Road to resilience: a systematic review and meta-analysis of resilience training programmes and interventions." *BMJ Open* 8(6): e017858.
- Loprinzi CE, Prasad K, Schroeder DR, et al. 2011. "Stress Management and Resilience Training (SMART) program to decrease stress and enhance resilience among breast cancer survivors: a pilot randomized clinical trial." *Clinical Breast Cancer* 11(6): 364-368.
- Luthar SS, Cicchetti D, & Becker B. 2000. "The construct of resilience: A critical evaluation and guidelines for future work." *Child Development* 71: 543-562.
- Sood A, Prasad K, Schroeder D, et al. 2011. "Stress management and resilience training among Department of Medicine faculty: a pilot randomized clinical trial." *Journal of general internal medicine* 26.
- Steinhardt M and Dolbier C. 2008. "Evaluation of a resilience intervention to enhance coping strategies and protective factors and decrease symptomatology." *Journal of American College Health* 56(4): 445-453.
- Waldinger R & Schulz M. 2023. "The Good Life: Lessons from the World's Longest Scientific Study of Happiness"



## 2. | Are screens making young people sad and anxious?



Technology use among young people is growing increasingly ubiquitous and a key part of their lives. Recent estimates suggest that teenagers spend an average of 30 hours in front of screens each week<sup>7</sup>. At the same time, mental health problems and psychological distress are increasing in young people, from 26% to 40% over the past two decades<sup>8</sup>. A common narrative is that technology use is a key driver of the youth mental health crisis. However, the reality is that the science and knowledge to definitively understand the effects of technology use on young people’s mental health is lacking. Understanding this issue is one of the biggest and most critical challenges in addressing mental health and wellbeing.

## Correlation or causation?

Existing literature has established a correlational link between screen time and mental health problems including depression and anxiety in young people, showing that the two are related in a linear fashion. For example, findings from the Future Proofing Study, which is a large longitudinal study of adolescent mental health led by the Black Dog Institute, found a significant relationship between screen time and symptoms of depression, especially among young girls<sup>9</sup>.

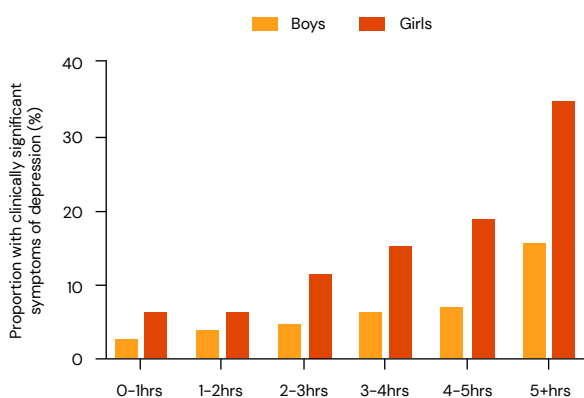


Figure 1: Relationship between screen time and clinically significant symptoms of depression. Black Dog Institute (2022)

While screen time and depression demonstrate a strong association, the direction of this relationship is unclear. It is widely assumed that technology use leads to mental health problems, although the possibility remains that experiencing mental health problems such as depression, lead to young people increasingly turning to screens. Longitudinal data in this area is lacking, and of what exists, the direction of the relationships is not clear or reliable<sup>10</sup>, and when effects exist, they are small in magnitude<sup>11</sup>. Overall, there is little consensus or evidence that screentime negatively impacts mental health.

## How young people use screens

Not all screen time and online activities are equal in nature. It is important to understand the differences in what young people are doing online.

For example, research has found that children who actively connect and communicate with friends online have larger peer groups, while passive use of screens such as extensive video streaming is associated with mental health symptoms<sup>12</sup>.

7 Thomas et al., 2020

8 Black Dog Institute, 2022

9 Ibid.

10 Tang et al., 2021

11 Orben & Przybylski, 2019

12 Paulich et al., 2021

Moreover, there is an emerging field in the area of digital emotion regulation which examines the way in which people are using online technologies to regulate and shape their emotional states<sup>13</sup>.

Increasingly, the literature is leading towards the idea that it is how young people (and indeed, people in general) are spending their time online and their patterns of technology use that is important in determining whether it is associated with benefits or harm.

### **Lost time**

The displacement hypothesis is the idea that technology use might be displacing other important activities for health and wellbeing, such as sleep, physical activity and in-person interaction, and it is for this reason there may be a negative impact of technology use. For example, there is strong evidence that device use before bed delays sleep among teenagers<sup>14</sup>.

### **From research to policy**

Given the complexities and gaps in knowledge about how technology use, mental health and wellbeing are associated, it is not surprising that there is confusion among policy makers, schools, families and young people themselves about how best to engage with the world online.

This confusion is evidenced by inconsistent policies implemented across the country, such as school phone bans in some states but not others. Further, current guidelines about screen use advise no more than two hours of use for

5–17-year-olds<sup>15</sup>. These guidelines currently do not reflect the reality, and more nuance around helpful compared to less helpful uses of technology are needed. More specifically, policy frameworks that include clear guidelines for age-appropriate technology use, and education that teaches young people and their parents about the benefits and risks of screen time are required.

### **Evidence-based interventions**

Digital literacy programs delivered in schools and the community tend to focus on risks and safety in navigating the online world, and there is some outstanding work in the space being led by the eSafety Commission.

What is lacking is education about what healthy patterns of technology use looks like in everyday life. We believe that we have a responsibility to prioritise using high quality longitudinal data to answer this question about healthy technology use, and then use this information to upskill schools and parents to provide guidance to young people about responsible screen use.

This should be developed collaboratively with young people, parents, researchers and health professionals to provide young people and their support networks with strategies to navigate the digital world.

Technology and smartphones have become an integral part of life in the modern world. We need to learn more about the complex relationship between screens and mental health, and then use these learnings to upskill the community in

---

13 Wadley et al., 2020

14 Hale & Guan, 2015

15 AIFS, 2021

digital literacy so that we can guide young people towards a balanced and healthy relationship with screens.



### Discussion questions

1. Why do you think technology use might be harmful to young people? What are you seeing in your community?
2. What aspects of technology use might be useful to young people? What are you seeing in your community?
3. What actions can industry and government make to help young people use technology in a positive way?
4. What challenges do you anticipate in this area in the future?
5. What questions would you like answered by researchers?

### References

- Australian Institute of Family Studies. (2021). Too much time on screens? Screen time effects and guidelines for children and young people. Accessed: <https://aifs.gov.au/resources/short-articles/too-much-time-screens>
- Black Dog Institute. 2022. "Turning the Tide on Depression." White Paper.
- Hale, L., & Guan, S. 2015. "Screen time and sleep among school-aged children and adolescents: a systematic literature review." *Sleep Med Rev* 21: 50–58. <https://doi.org/10.1016/j.smr.2014.07.007>
- Hale, L., Kirschen, G. W., LeBourgeois, M. K., Gradisar, M., Garrison, M. M., Montgomery-Downs, H., Kirschen, H., McHale, S. M., Chang, A. M., & Buxton, O. M. 2018. "Youth Screen Media Habits and Sleep: Sleep-Friendly Screen Behavior Recommendations for Clinicians, Educators, and Parents." *Child Adolesc Psychiatr Clin N Am* 27 (2): 229–245. <https://doi.org/10.1016/j.chc.2017.11.014>
- Orben, A., & Przybylski, A. K. 2019. "The association between adolescent well-being and digital technology use." *Nature Human Behaviour* 3 (2): 173–182. <https://doi.org/10.1038/s41562-018-0506-1>
- Paulich, K. N., Ross, J. M., Lessem, J. M., & Hewitt, J. K. 2021. "Screen time and early adolescent mental health, academic, and social outcomes in 9- and 10-year-old children: Utilizing the Adolescent Brain Cognitive Development (ABCD) Study." *PLoS One* 16 (9): e0256591. <https://doi.org/10.1371/journal.pone.0256591>
- Sewall, C. J. R., Goldstein, T. R., & Rosen, D. 2021. "Objectively measured digital technology use during the COVID-19 pandemic: Impact on depression, anxiety, and suicidal ideation among young adults." *Journal of Affective Disorders* 288: 145–147. <https://doi.org/10.1016/j.jad.2021.04.008>
- Tang, S., Werner-Seidler, A., Torok, M., Mackinnon, A. J., & Christensen, H. 2021. "The relationship between screen time and mental health in young people: A systematic review of longitudinal studies." *Clin Psychol Rev* 86: 102021. <https://doi.org/10.1016/j.cpr.2021.102021>
- Thomas, G., Bennie, J. A., De Cocker, K., Ireland, M. J., & Biddle, S. J. H. 2020. "Screen-based behaviors in Australian adolescents: Longitudinal trends from a 4-year follow-up study." *Preventive Medicine* 141: 106258.
- Wadley, G., Smith, W., Koval, P., & Gross, J. J. 2020. "Digital Emotion Regulation." *Current Directions in Psychological Science* 29 (4): 412–418. <https://doi.org/10.1177/0963721420920592>



### 3. | Blended care: combining human and digital in clinical care

There is a significant gap between the number of people who need mental care and those receiving it in Australia<sup>16</sup>. Over 20 years of research has shown that digital therapies are clinically effective, cost-efficient and are a viable option to improve access, equity and quality of care, especially for hard-to-reach populations. However, these tools can suffer from lower engagement and high drop-out when people try them as standalone interventions, without the support of a clinician. A blended model of care that combines the best of both digital technology and the skill and training of a mental health professional delivering face-to-face treatment is a potential solution to closing the mental health treatment gap in Australia.

## Mental health treatment gap

Less than half of people who need mental health treatment are accessing it<sup>17</sup>. Of those that do, only 14% receive a full dose of evidence-based treatment<sup>18</sup>. The situation is far worse for Australians living in regional, rural and remote areas<sup>19</sup>. Consumers frequently report that the mental health system in Australia is confusing, fragmented, and complex to navigate. There are several significant barriers to accessing timely and quality face-to-face mental health treatment, including an overstretched workforce, leading to long waiting times and high gap payments that are unaffordable for many<sup>20</sup>.

Last year, 75% of psychologists had a wait list and one third were unable to take on new patients. For people who are already in care, this shortage also impedes mental health professionals' ability to provide services that are high quality and patient-centred. Medicare rebates are insufficient in

scope (10 session/year) to provide an evidence-based treatment, and high gap payments (>\$100/session) place high financial burden on patients, rendering treatment often unaffordable to many on low to middle incomes<sup>21</sup>.

## Digital solutions

Digital interventions for depression and anxiety have been developed, tested, and disseminated to overcome barriers to care (e.g., geography, lack of workforce, costs). Digital interventions can provide similar treatment outcomes to in-person care but are more cost-effective as they can be delivered with 10% of clinician time required by in-person care<sup>22</sup>. These therapies improve accessibility, quality and standardisation of care, and several excellent examples exist in Australia and already form an important part of the mental health system.

16 Islam et al., 2022; Petrie et al., 2021; Productivity Commission, 2020; State of Victoria, 2019

17 Whitehall et al., 2014

18 Harris et al., 2015

19 Spijker et al., 2019; Meadows et al., 2015

20 Petrie et al., 2021; Academic Unit of General Practice [AUGP], 2020; Dulsen et al., 2020

21 Pirkis, J., Currier, D., Harris, M & Mihalopoulos, C., 2022

22 Carlbring et al., 2018; Andrews et al., 2018; Weisel et al., 2019; Karyotaki et al., 2019

## Limitations of digital

However, digital services are not typically integrated into the rest of the mental health system and often suffer from a drop in consumer engagement over time, meaning people do not receive treatment at the appropriate dosage.

Another barrier to greater uptake is a lack of trust among consumer and clinicians due to low quality assurance of mental health apps. Today, tens of thousands of mental health apps are available on the app stores, however, studies show only 4% of these apps are based on quality evidence. The National Safety and Quality Digital Mental Health (NSQDMH) Standards provided accreditation for mental health apps but, unlike pharmaceuticals or medical devices, mental health apps are not required to be accredited to make claims about efficacy or safety.

## Blended care

A blended model of care that combines the best of both digital technology and the skill and training of a mental health professional delivering face-to-face treatment is a potential solution to closing the mental health treatment gap in Australia. In a blended care model, the digital component of treatment might support education, skills training, assessment and progress tracking, while the face-to-face (or telehealth) components are led by a trained mental health professional (e.g., psychologist) and focused on

more complex therapy work, such as boosting client engagement, and tailoring therapy skills to the client's unique personal situation. The combination of digital and face-to-face therapy can help improve patient motivation and create longer term behavioural change.

Evidence from recent European studies show that blended care, which combines digital and face-to-face psychological therapy, is more effective at reducing depression, and reduces treatment length and costs compared to face-to-face therapy alone<sup>23</sup>. Integrating digital services into the current mental health treatment system can help to put the person at the centre of their care, increasing access and consumer choice.

## Increasing trust and uptake

Mental health reform will take years. In the meantime, there are some actions we can take to facilitate increased trust and uptake of blended care models. Training clinicians in digital and blended models of care will be fundamental. It is also critical to look at reimbursement models as part of the Medicare review and consider a new MBS item number that would incentivise clinicians to prescribe, educate and monitor patients on digital mental health treatments. The development of digital infrastructure is also critical to ensure consumers' mental health data is secure and their privacy protected as these services will require additional data collection and storage.

---

23 Etzelmueller, C. et al 2020.



## Discussion questions

1. What is the role of digital in closing the mental health service gap?
2. Why are digital technologies not better integrated into the mental health system?
3. How can we increase trust and uptake of digital and blended mental health care models?
4. What is the role of government and industry in assuring quality and safety in digital mental health products?
5. How do you see digital mental health tools connecting with electronic health records in a way that increases continuity of care while protecting consumers' privacy?

## References

- Andrews G, Basu A, Cuijpers P, et al. 2018. "Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis." *J Anxiety Disord* 55: 70–78.
- Andrews, G., Newby, J. M., & Williams, A. D. 2015. "Internet-delivered cognitive behavior therapy for anxiety disorders is here to stay." *Current psychiatry reports* 17(1): 533. <https://doi.org/10.1007/s11920-014-0533-1>
- Bregje A van Spijker, Jose A Salinas-Perez, [...], and Luis Salvador-Carulla. 2019. "Service availability and capacity in rural mental health in Australia: Analysing gaps using an Integrated Mental Health Atlas". Volume 53, Issue 10. <https://doi.org/10.1177/0004867419857809>
- Deady, M., Tan, L., Kugenthiran, N., Collins, D., Christensen, H., Harvey, S. B. 2020. "Unemployment, suicide and COVID-19: using the evidence to plan for prevention."
- Dulsen, P., Bendig, E., Kuchler, AM., Christensen, H., Baumeister, H., (2020) Digital interventions in adult mental healthcare settings: recent evidence and future directions. *Co-Psychiatry*. Op Ed.
- Etzelmueller, C. Vis, E. Karyotaki, H. Baumeister, N. Titov, M. Berking, P. Cuijpers, H. Riper, D.D. Ebert. 2020. "Effects of internet-based cognitive behavioral therapy in routine care for adults in treatment for depression and anxiety: systematic review and meta-analysis." *J. Med. Internet Res.* 22, Article e18100.
- Karyotaki E, Riper H, Twisk J, et al. 2017. "Efficacy of Selfguided Internet-Based Cognitive Behavioral Therapy in the Treatment of Depressive Symptoms: A Meta-analysis of Individual Participant Data." *JAMA Psychiatry* 74(4): 351-359.
- Pirkis, J., Currier, D., Harris, M & Mihalopoulos, C. (2022) 'Evaluation of Better Access'. University of Melbourne. <https://www.health.gov.au/sites/default/files/2022-12/main-report-evaluation-of-the-better-access-initiative.pdf>
- Productivity Commission. 2019. *Mental Health Inquiry Draft Report, Chapter 6 Supported Online Treatment*. Australian Government, Canberra.
- State of Victoria. 2019. *Royal Commission into Victoria's Mental Health System, Interim Report, Parl Paper No. 87 (2018–19), Part 5, 20.2 Need for increased investment*. Victorian Government.
- Weisel, K.K., Fuhrmann, L.M., Berking, M. et al. 2019. "Standalone smartphone apps for mental health— a systematic review and meta-analysis." *npj Digit. Med.* 2, 118. <https://doi.org/10.1038/s41746-019-0188-8>
- Whiteford, H.A., et al. 2014. "Estimating treatment rates for mental disorders in Australia." *Aust Health Rev* 38(1): p. 80–5.



## 4. | Rise of AI in mental health - risks and opportunities



Artificial intelligence (AI) has exploded in the past year. Products like ChatGPT – which answers complex questions – and Midjourney – which creates vivid imagery – are so human-like that they have captured the world’s imagination. AI excitement has found its way to mental health, where scientists and developers are racing to create AI-based tools that can assess, diagnose, and treat mental illness at a previously unimaginable scale. Early efforts are both impressive and promising, but with them come real risks of harm. A measured approach will be needed to create AI tools that are safe and effective for the vulnerable millions across the world who urgently need mental health care.

## What’s new about AI?

AI shares some similarities with human intelligence but the two are very different. AI models like the Generative Pre-trained Transformers that give ChatGPT its name borrow from our brain’s extraordinary capacity to create predictions based on things we have learned. These AI models can scan an enormous amount of data very quickly and use this ‘knowledge’ to create something new based on requests from humans like, “Write me a treatment plan for my depression”. These generative models and other forms of AI are poised to change the way we diagnose and treat mental illness – hopefully for the better, but possibly for the worse.

## Moving beyond diagnosis

Mental illness looks different to every individual. Some treatments are one-size-fits all; others use a trial-and-error approach that is not always based on evidence. The Diagnostic and Statistical Manual of Mental Disorders (DSM) has long formed the basis how mental illness is diagnosed. But it, too, does not reflect our best evidence.

Happily, science is moving away from a bento-box view of mental illness, where diagnoses are categorical sets of symptoms, to a more transdiagnostic approach, where treatments are personalised to individual needs. Neuroscience is increasingly using a form of AI called ‘machine learning’ to scan massive amounts of brain data and decipher how someone’s brain produces their thoughts, emotions, and behaviours. This could lead to a paradigm shift in psychiatry where mental illnesses are treated based on underlying physiology, rather than what someone’s symptoms ‘look like’ to a doctor<sup>24</sup>.

## Precision mental health

Early evidence shows that ‘precision medicine’ approaches offered by machine learning can be more effective than a typical digital health treatment for depression. Beyond machine learning, natural language processing (NLP) is also gaining significant momentum in mental health research. NLP models can now scan recordings and notes from therapy sessions to give therapists insights into what is working and when<sup>25</sup>, allowing therapists to finetune

24 Chen et al., 2022  
25 Burger et al., 2021

their therapeutic approach to each client. For example, mental health start-up Kintsugi is using NLP models to identify “voice biomarkers” that reportedly analyse phone calls to healthcare providers to predict a person’s symptoms of depression and or anxiety, alerting non-mental health practitioners that a patient may be suffering psychological distress. While this may be useful in detecting unseen mental health concerns, some users may prefer to not disclose mental health matters to certain providers. Patients in distress may be reluctant to seek help at all if they know this technology is being used. And what happens when the AI model gets it wrong? There are a few potential pitfalls of AI that are worth bearing in mind as we adopt the technology in mental health care.

### **When AI gets it wrong**

Generative AIs, like ChatGPT, are not always as accurate as they seem. This may be fine when writing an email but could have serious consequences when writing a treating plan. Like humans, generative AIs can “hallucinate”, a term used to describe times when an AI gives a response that seems at odds with its training data. Rather than drawing on facts, these hallucinations are convenient fictions that fill space in the AI’s response. They may appear convincing, but they are completely fabricated. A related concept is a ‘stochastic parrot’, which refers to an AI that can generate convincing language but does not understand what that language means, so it doesn’t know if what it’s saying is wrong or inappropriate. AIs are only repeating what they learn from their training

data, so serious inaccuracies could be naively promoted by the AI and go undetected by a patient. The results could be AI therapists that fill gaps in their knowledge with convincing lies, all the while unaware of the danger they pose to their patients.

### **Ensuring quality and safety**

Digital mental health is plagued by products that claim to work but struggle to provide any convincing evidence or only spruik convincing-sounding evidence. Sadly, the AI space is no different. Some AI-based mental health start-ups claim their treatments work better than traditional approaches to care without providing proof. Research often focuses only on people with mild symptoms, and several prominent AI-based mental health start-ups provide enthusiastic summaries of small, low-quality trials<sup>26</sup>. There is significant work to be done to protect vulnerable consumers and support mental health practitioners to select safe, effective AI-based therapeutics.

Fortunately, this work is well underway. Organisations such as The Gradient Institute, are focusing their considerable efforts on understanding what ethical AI looks like and how it can be applied in high-risk areas, like mental health. In 2021, the World Economic Forum released their Global Governance Toolkit for Digital Mental Health<sup>27</sup>. In it they lay out principles for implementing AI, “with a community through transparency and shared decisions on the best course of actions”. These principles focus on realising the human benefits that AI promises in

---

<sup>26</sup> Inkster B, Sarda S, Subramanian V, 2018

<sup>27</sup> Allen, S., Bernaert, A., et al., 2021

a way that is not only safe and reliable, but also transparent and accountable. Through these principles we will likely find our way to a future where AIs transform mental health care and deliver effective mental health treatments across the globe.

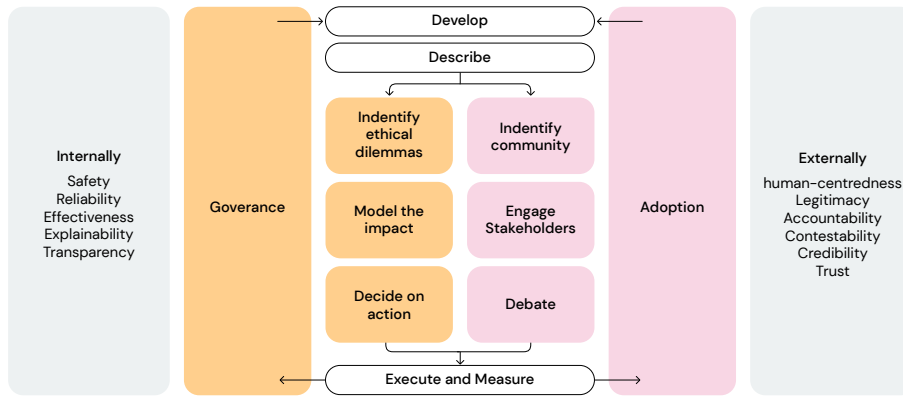


Figure 2: A framework for the ethical operation of AI in mental health. Allen, S., Bernaert, A., et al. (2021)



## Discussion questions

1. What uses for AI in mental health do you find most promising? Or most frightening?
2. How can we ensure AI mental health tools developed to a high-quality of accuracy and safety?
3. How can we develop mental health datasets that are privacy-focused and have limited biases?
4. How do we ensure AI mental health tools are culturally appropriate, especially for First Nations people and people from culturally and linguistically diverse communities?
5. How should we best regulate AI mental health tools?

## References

Allen, S., Bernaert, A., et al. 2021. "Global Governance Toolkit for Digital Mental Health: Building Trust in Disruptive Technology for Mental Health." World Economic Forum. [https://www3.weforum.org/docs/WEF\\_Global\\_Governance\\_Toolkit\\_for\\_Digital\\_Mental\\_Health\\_2021.pdf](https://www3.weforum.org/docs/WEF_Global_Governance_Toolkit_for_Digital_Mental_Health_2021.pdf)

Burger F, Neerincx MA, Brinkman W-P. 2021. "Natural language processing for cognitive therapy: Extracting schemas from thought records." PLoS ONE 16 (10): e0257832. <https://doi.org/10.1371/journal.pone.0257832>

Inkster B, Sarda S, Subramanian V. 2018. "An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study." JMIR Mhealth Uhealth 6 (11): e12106. doi: 10.2196/12106

Zhe Sage Chen, Prathamesh (Param) Kulkarni, Isaac R. Galatzer-Levy, Benedetta Bigio, Carla Nasca, Yu Zhang. 2022. "Modern views of machine learning for precision psychiatry." Patterns, Volume 3, Issue 11. doi.org/10.1016/j.patter.2022.100602



October 2023

**Science.**  
**Compassion.**  
**Action.**