

<https://doi.org/10.1038/s44184-024-00071-0>

Dynamic learning of individual-level suicidal ideation trajectories to enhance mental health care



Mathew Varidel¹ ✉, Ian B. Hickie¹, Ante Prodan^{1,2,3}, Adam Skinner¹, Roman Marchant^{4,5}, Sally Cripps^{4,5}, Rafael Oliveria⁶, Min K. Chong¹, Elizabeth Scott¹, Jan Scott⁷ & Frank Iorfino¹

There has recently been an increase in ongoing patient-report routine outcome monitoring for individuals within clinical care, which has corresponded to increased longitudinal information about an individual. However, many models that are aimed at clinical practice have difficulty fully incorporating this information. This is in part due to the difficulty in dealing with the irregularly time-spaced observations that are common in clinical data. Consequently, we built individual-level continuous-time trajectory models of suicidal ideation for a clinical population ($N = 585$) with data collected via a digital platform. We demonstrate how such models predict an individual's level and variability of future suicide ideation, with implications for the frequency that individuals may need to be observed. These individual-level predictions provide a more personalised understanding than other predictive methods and have implications for enhanced measurement-based care.

Routine outcome monitoring has been promoted to enhance mental health care. The aim is to record the longitudinal characteristics of an individual's mental health using digital tools, such as websites and apps with patient-report outcome measures^{1–6}, and passive monitoring or wearable devices^{7,8}.

Within the clinical setting, this routine assessment, typically termed measurement-based care (MBC^{9–13}), provides more objective insights into a person's condition, treatment response and progress. This approach improves clinical decision-making about treatment adjustments, ensuring that interventions are tailored to a person's needs¹⁴. It has been shown to improve outcomes, retention, monitoring and clinician-individual communication^{15–17}. As such, MBC has often been recommended for use in clinical services, and several digital technologies have been developed to help implement it (e.g. Innowell³, NovoPsych and Greenspace).

While MBC enhances the quality of care, interpreting measurements within an individual's context for personalised decisions is crucial. Clinicians require predictive capability as interventions are made with the goal of improving an individual's future mental health and functioning¹⁸. Unfortunately, the predictive performance of clinicians tends to be poor^{19–21}, emphasising the need for tools to improve this capability.

The use of predictive models can provide insight into an individual's likelihood of a specific outcome, although their compatibility with the ongoing and personalised nature of MBC is less clear^{22,23}. Typically, they predict outcomes over defined periods, and they may not update dynamically with new observations. An alternative approach is to analyse and predict trajectories of outcomes over time, offering insight into an individual's progress for ongoing decision-making^{24–26}. However, current implementations are limited as they make predictions from a baseline or current observation and typically fail to incorporate historical observations, which hinders their practical use in real-world settings.

Categorical predictions for suicidal thoughts and behaviours are challenging, with models only performing slightly better than chance with respect to positive predictive value^{27–29}. While a number of factors have been shown to be associated with suicidal thoughts and behaviours including: prior suicidal thoughts and behaviours, clinical diagnoses and specific symptoms^{28,30}, functioning^{31,32} and socio-demographic factors³³, this has not translated to high positive predictive values. To improve our understanding and prediction there is increased focus on higher frequency (e.g. hours, days or weeks) longitudinal observations. These studies have found that suicidal thoughts fluctuate over short time intervals with differences across groups of individuals^{34,35}, and due to a range of factors³⁶ including negative mood,

¹Brain and Mind Centre, The University of Sydney, Sydney, NSW, Australia. ²Translational Health Research Institute, Western Sydney University, Sydney, NSW, Australia.

³School of Computer, Data and Mathematical Sciences, Western Sydney University, Sydney, NSW, Australia. ⁴Human Technology Institute, University of Technology, Sydney, NSW, Australia. ⁵School of Mathematical and Physical Sciences, University of Technology Sydney, Sydney, NSW, Australia. ⁶Data61, CSIRO, Sydney, NSW, Australia. ⁷Academic Psychiatry, Institute of Neuroscience, Newcastle University, Newcastle, UK. ✉e-mail: mathew.varidel@sydney.edu.au

hopelessness, loneliness, affective instability and poor sleep^{37,38}. Recognising this fluctuation and understanding the range of possible trajectories for an individual is crucial given the inherent challenges of prediction.

Considering these issues, we advocate for a continuous time modelling framework to supplement MBC³⁹. Continuous time models have several properties that make them appealing for clinically collected data including the ability to: (1) deal with irregularly time-spaced intervals and missing domains^{40,41}; (2) be framed in a hierarchical manner such that parameters can be learned for the individual but constrained by the population-level^{42,43}; (3) be framed in a Bayesian framework to understand uncertainties and thus the range of possible outcomes, and (4) have their underlying state values transformed onto the typically bounded and discretised (e.g. Likert) scales via link functions. Continuous time models have been applied in many fields (e.g. finance for 50 years⁴⁴), yet their application in mental health is limited with discrete time models as the norm for frequent data sampling methods. We suspect that the properties of continuous time models make them particularly appropriate for modelling naturalistically collected clinical data within an MBC framework, which to our knowledge, has not yet been explored.

In this paper, we explore the use of continuous time models for naturally collected data in a clinical setting. We focus solely on building and comparing several model parameterisations for longitudinal data on suicidal ideation. After selecting the most appropriate model, we then show specific insights that can be learned from these models that could aid clinical decision-making.

Results

Sample characteristics

The sample consisted of 585 people, 72.6% female, with a mean age of 24.2 years (SD, 10.8 years). Of those included in the analysis, 238 (40.7%) had prior suicidal plans or attempts within their lifetime and 61 (10.4%) reported plans or attempts within the last three months. The mean number of observations for suicidal ideation per individual was 4.5, although it was skewed to smaller numbers with a median of 2, and more specifically 55 people with at least 10 and 140 people with at least 5 observations. The time difference between observations had a median of 35 days (IQR, 1–108 days). Other baseline characteristics of the sample are shown in Table 1.

Model comparison

The Wiener process (i.e. continuous time random walk process) with random effects for all parameters was the best predictive model (see Table 2). Widely applicable information criteria (WAIC^{45–47}) estimates were calculated as a measure of model performance using a subset of the data for individuals with ≥ 10 observations ($N = 55$). Various model formulations were explored by allowing random effects for different parameters. Initially, we assumed a Wiener process with mixed effects (random baseline but fixed diffusion). The Wiener process with random effects performed much better, which suggested that the magnitude of diffusion is different across individuals. The Ornstein–Uhlenbeck process (i.e. continuous time autoregressive process) adds an extra effect that drives individuals towards a constant value over time. We use the Wiener process models with random effects for all further analyses in this paper.

Population and individualised parameter distributions

Our final model has two key parameters that represent a transformed baseline observation ($\phi_{m,1}$) and a diffusion over time parameter ($\phi_{m,2}$) for each individual m . We estimated a full posterior distribution for both variables for each individual. In Fig. 1 we show the median of the diffusion parameter and the predicted baseline observation, which is a transformation of the baseline parameter to the suicide ideation scale. The marginal histograms show the range of values that an individual could have given the sampled population parameters (see Supplementary Methods 5 for further detail). We see no clear dependency between these two parameters. Although, there is a large group of individuals that have zero suicidal ideation at baseline.

The diffusion is an unbounded model parameter where higher values correspond to greater observed variability in the trajectory. The values are somewhat abstract given their non-linear mapping to the observational space, but as an example, assuming a starting suicidal ideation score of 2 and the median population diffusion parameter of -1.23 a trajectory will diffuse to a range of [1, 4] within a week.

Table 1 | Baseline characteristics of the study participants used for modelling suicidal ideation

	Total	Female	Male
No. (%)	585	425 (72.6%)	160 (27.4%)
Demographics			
Age, mean (SD)	24.2 (10.8)	23.4 (9.9)	26.6 (12.8)
Urban ^a , No. (%)	503 (86.0%)	364 (85.6%)	139 (86.9%)
Health history, No. (%)			
Any prior mental health problem	399 (68.2%)	289 (68.0%)	110 (68.8%)
Any prior mental health treatment	402 (68.7%)	287 (67.5%)	115 (71.9%)
Any traumatic event	221 (37.8%)	164 (38.6%)	57 (35.6%)
Any physical health problem	175 (29.9%)	132 (31.1%)	43 (26.9%)
Any family mental health problem	316 (54.0%)	235 (55.3%)	81 (50.6%)
Prior suicidal plans or attempts (CSSRS, Question 3)	238 (40.7%)	169 (39.8%)	43 (26.9%)
Three-month prior suicidal plans or attempts (CSSRS, Question 4)	61 (10.4%)	47 (11.1%)	14 (8.8%)
Mental health, median (Q1, Q3)			
Distress (K10)	31 (25, 36)	32 (26, 37)	30 (23, 36)
Depressed mood (QIDS)	14 (10, 18)	15 (11, 18)	13 (9, 16)
Anxiety (OASIS)	10 (6, 13)	10 (6, 13)	10 (6, 12)
Psychosis-like experiences (PQ16)	4 (1, 7)	4 (1, 7)	4 (2, 7)
Manic-like experiences (ASRM)	2 (0, 4)	2 (0, 4)	3 (1, 5)
Suicidal ideation (SIDAS)	5 (0, 17)	5 (0, 17)	5 (0, 14)
Social support (Schuster's SSS)	7 (5, 9)	7 (4, 9)	6 (5, 8)

^aLocations could not be determined for 39 individuals.

Table 2 | Comparison of predictive performance

Number	Type	Effects				WAIC
		Baseline ($\phi_{.1}$)	Diffusion ($\phi_{.2}$)	Drift ($\phi_{.3}$)	Constant ($\phi_{.4}$)	
4	Ornstein–Uhlenbeck	Random	Random	Random	Random	5390.4
3		Random	Random	Fixed	Random	5453.4
2	Wiener	Random	Random	NA	NA	5303.8
1		Random	Fixed	NA	NA	5460.2

The first two columns provide a model descriptor with a number that increases with the number of free parameters and model type. The next four columns outline whether the parameters were random or fixed. The parameters define the initial state ($\phi_{.1}$), diffusion ($\phi_{.2}$), lagged autoregressive effect ($\phi_{.3}$) and a long-term constant ($\phi_{.4}$). Further detail regarding these parameters can be found in Supplementary Methods 2. The final column provides a measure of model performance estimated using the WAIC.

We also check for correlations between the diffusion parameter and the sample characteristics at baseline (see Table 3). Here, we find that the diffusion parameter is positively correlated with the last three-month history of suicide attempts or planning (mean, 0.10, ETI [95% equal-tailed interval], 0.02, 0.19) and borderline correlated with mania-like experiences (mean, 0.07, ETI, 0.00, 0.14).

Individual trajectory learning and prediction

Our focus was on constructing individualised predictions. As such, we show several illustrative examples of individuals within our data. We show six

Table 3 | Correlation mean (95% ETI) between baseline characteristics and the diffusion parameter

Demographics	Correlation
Age	−0.01 (−0.08, 0.07)
Sex (female–male)	−0.19 (−0.53, 0.15)
Urban (urban–regional)	−0.17 (−0.40, 0.05)
Health history	
Any prior mental health problem	0.01 (−0.06, 0.08)
Any prior mental health treatment	−0.01 (−0.08, 0.07)
Any traumatic event	0.02 (−0.06, 0.09)
Any physical health problem	−0.01 (−0.08, 0.07)
Any family mental health problem	0.00 (−0.08, 0.07)
Prior suicidal plans or attempts (CSSRS, Question 3)	−0.01 (−0.08, 0.06)
Three-month prior suicidal plans or attempts (CSSRS, Question 4)	0.10 (0.02, 0.19)
Mental Health	
Distress (K10)	0.02 (−0.05, 0.09)
Depressed mood (QIDS)	0.01 (−0.07, 0.08)
Anxiety (OASIS)	0.03 (−0.04, 0.11)
Psychosis-like experiences (PQ16)	0.04 (−0.04, 0.11)
Manic-like experiences (ASRM)	0.07 (0.00, 0.14)
Suicidal ideation (SIDAS)	0.01 (−0.05, 0.07)
Social support (Schuster's SSS)	0.06 (−0.01, 0.12)

For the binary variables, we calculate the mean (95% ETI) for the difference in the diffusion parameter for each group.

individuals with qualitatively different historical trajectories that are broadly representative of the diversity of trajectories across the population, along with their predicted trajectories from their final observation in Fig. 2. We include historical trajectories that deteriorated (a) and improved (b), as well as those with moderate ideation (c), high variability (d) or moderate variability but with high ideation (f). We also include an individual who had no suicidal ideation while in care (e), which accounts for 152 (26.0%) individuals within our population.

We then show the dynamic process by which predictions are updated for an individual as observations occur. This is an approximation for the process by which individual trajectories would be estimated within a real-world clinical setting as data is collected. This is shown in Fig. 3. We show an individual that has moderate variability in suicidal ideation throughout care. The possible range of trajectories is highly uncertain at baseline (Fig. 3a) but as observations are collected over time (Fig. 3b–d) the ETI ranges decrease, suggesting that the model is learning that this individual has relatively stable suicidal ideation over time.

Prediction summaries

For each set of predictions, we provide several measures to quantify the level and variability of future suicidal ideation. This includes an estimate of the probability that an individual will be in the high-ideation category over the next 60 days. The variability measure V is an estimate of future observational variability with the distribution of median variability of 10.2 (IQR, 0.0–36.9) across individuals with ≥ 10 observations calculated from the final observation. We then provide a measure for how long it takes information to be 'lost' that we use as a recommended follow-up time (FUT), which has a median of 6 (IQR, 3–19 days) for individuals with ≥ 10 observations calculated from the final observation.

Discussion

We constructed continuous-time trajectory models for suicidal ideation using naturalistically collected data in a clinical setting. These dynamic models consider an individual's current and past observations and population parameters, to learn and update predictions as new observations are collected. We then provided the probability of an individual being in the high-ideation category over a particular time (i.e. IHIP) as well as an estimate for how long it takes to lose information (i.e. FUT). Several insights and applications can be drawn from these analyses.

The best-performing model was the Wiener process model which suggests that the typical trajectory is dominated by variability with no long-term constant. We note that this result will be partly due to

Fig. 1 | The joint distribution for the median values of the individual-level parameters. We show the median predicted baseline observation ($\tilde{y}_{m,0}$) compared to the diffusion parameter ($\phi_{m,2}$) for each individual m . This is compared to the prior predictive distribution for each parameter given the population-level parameters, where we show the median and 95% highest density credible intervals (HDI⁶⁶).

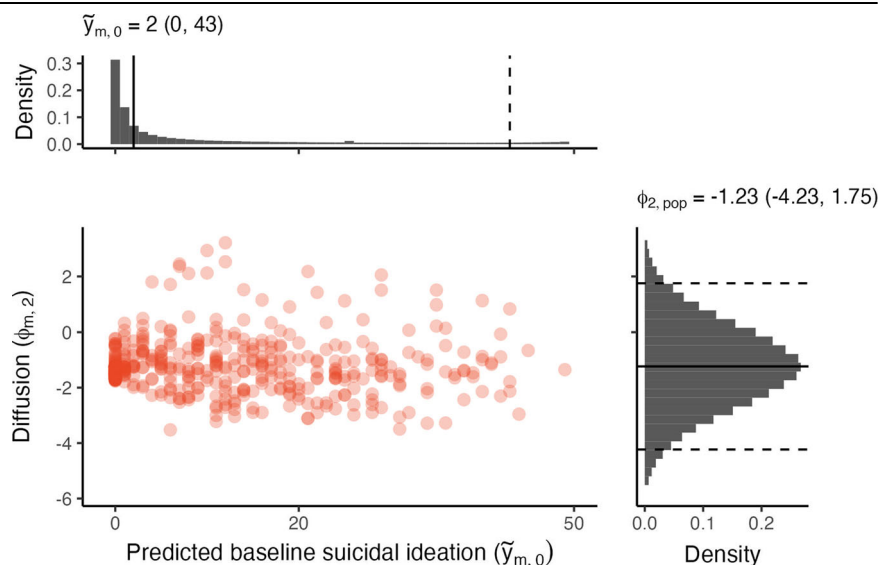
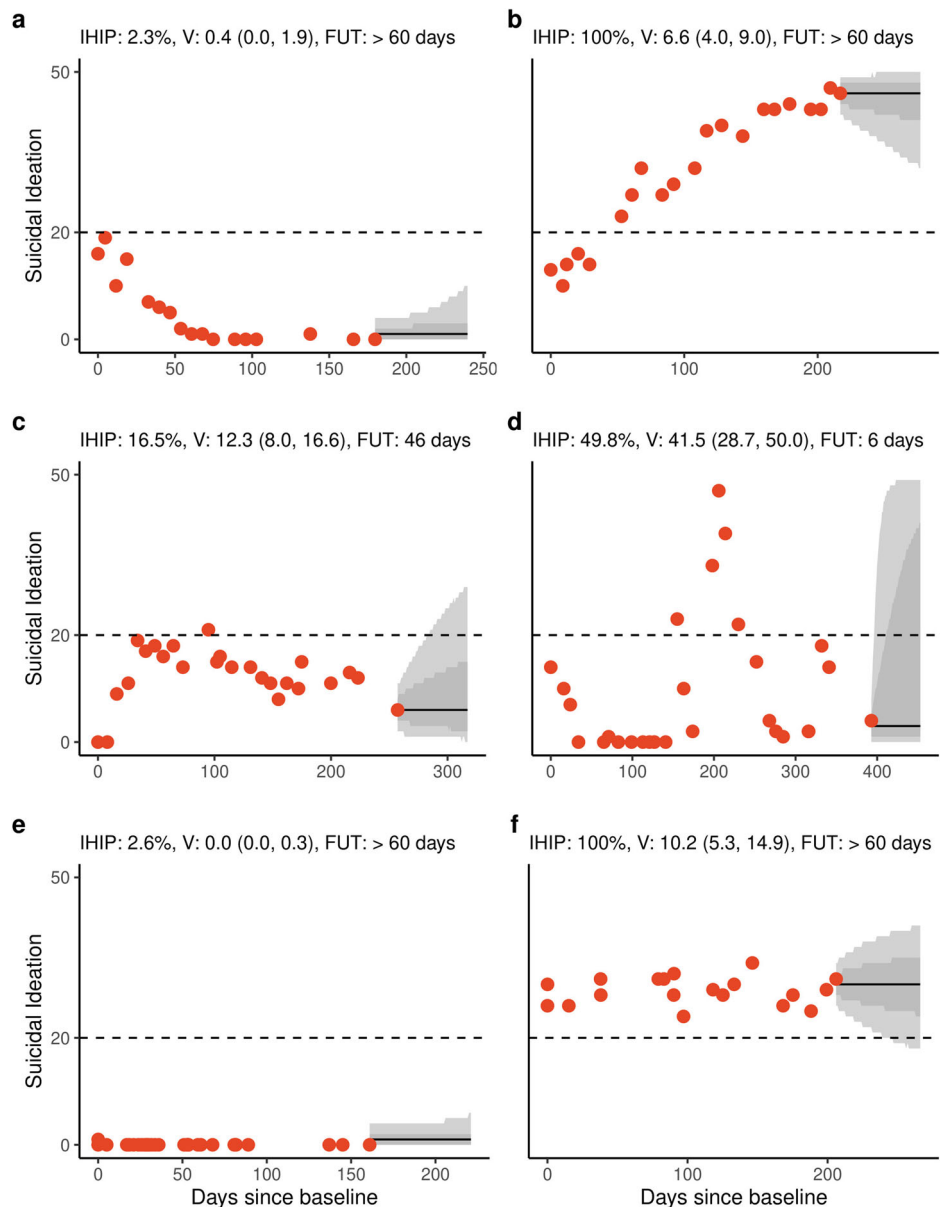


Fig. 2 | Predictive trajectories for individuals with qualitatively different historical trajectories. The distribution of predicted trajectories 60 days from the last observation for representative individuals that have **a** improved, **b** deteriorated, **c** had moderate risk and volatility, **d** high volatility, **e** no suicidal thoughts, and **f** high risk but moderate volatility. The distribution of trajectories is summarised using the median (black line), 68% (dark grey) and 95% (light grey) ETI each day. To summarise the individual's future ideation trajectories, we show the probability that the individual will be in the high-ideation category on any day during the next 60 days (integrated high-ideation probability, IHIP), the future observational variability (V) and a recommended follow-up time (FUT).



assuming that there is one type of model used across the population. Some individuals qualitatively appear to have a stable point, which suggests that a mixture of the two models would be ideal to identify those individuals that have a long-term constant that their trajectories trend towards.

Our work illustrates that there are qualitatively different trajectories of suicidal ideation among those in care. First, our population can be split into a group that had zero suicidal ideation throughout care (26.0%) and 'suicidal ideators' who have a wide range of potential trajectories that are poorly known at baseline. This complicates decision-making due to a poor understanding of an individual's future state, which means that decisions are typically reactive, as opposed to proactive. Despite the high degree of uncertainty about trajectories at baseline, individual trajectories can be learned and predictions can be dynamically updated as more information is collected, which could inform clinicians within MBC settings.

The use of trajectory modelling in MBC settings is still relatively novel, where current approaches tend to predict an expected trajectory by matching an individual's baseline or latest observations to a subset of the population^{25,26}. The comparison between this expected trajectory and an individual's realised trajectory is used to determine treatment progress and

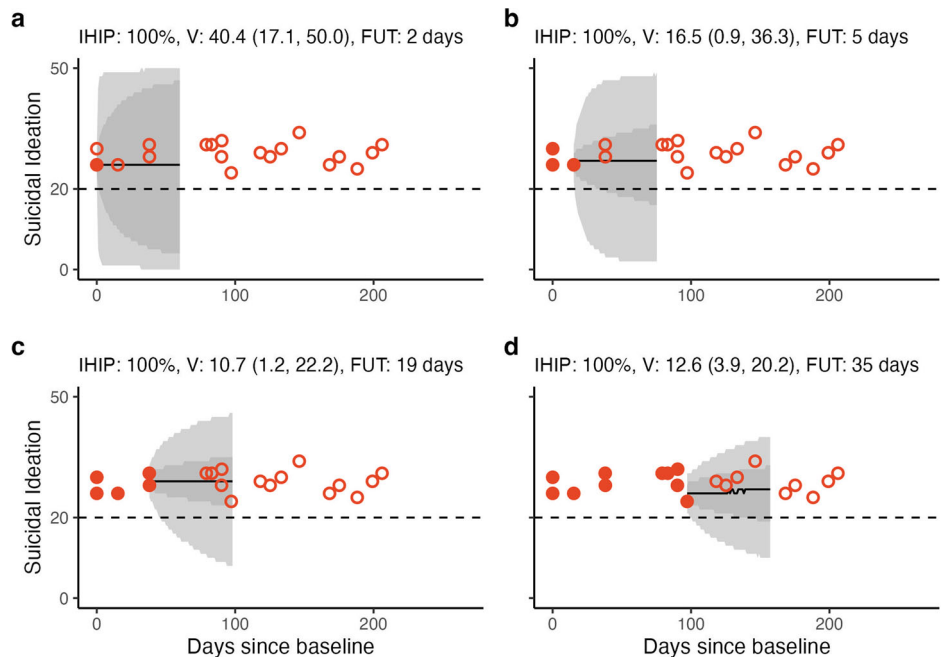
guide clinical decisions. While clearly useful as an approach for MBC, the reliance on an expected trajectory may limit the level of personalisation that could be achieved to effectively tailor interventions to a person's specific context and history.

An alternative or additional method to using an expected trajectory is to compare an individual's observations to their prior individual-level predictions. This would allow clinicians to understand when an individual's observations lie outside their own predicted range of trajectories (i.e. the 68% or 95% ETI). For example, individuals who have suicidal thoughts that emerge with no prior history throughout care will tend to lie outside of the predicted range of trajectories. This could be used as an additional flag for clinicians to understand this emergent phenomenon.

We found that suicidal ideation variability differs across our population. Our model comparison results showed that individualised diffusion parameters were required to model the data. We saw no clear demarcation within this distribution to suggest that there are categorical groups. Instead, we suspect that individuals lie on a continuum of variability, which emphasises that we should estimate trajectories at the individual level.

Differentiating the level of variability is important as it is suspected to lead to different outcomes⁴⁸. It has been argued that those with higher

Fig. 3 | Predictive trajectories for an individual as data is collected. Predictions for 60 days into the future from **a** baseline, and then after **b** three, **c** five, and **d** 10 observations. Each set of predictions uses the observations at or prior to that time (filled red circles) which are compared to future observations (red outlined circles). The distribution of trajectories is summarised using the median (black line), 68% (dark grey) and 95% (light grey) ETI for each day. To summarise an individual's future ideation trajectories, we show the probability that they will be in the high-ideation category on any day during the next 60 days (integrated high-ideation probability, IHIP), the future observational variability (V) and a recommended follow-up time (FUT).



variability have heightened stress responsivity compared to those with lower variability^{35,48}. The higher variability individuals may also be more likely to engage in spontaneous suicidal behaviours that would warrant an ongoing need to monitor their care closely and establish personal and social mechanisms to limit such precipitous behaviours. Whereas those that have high and persistent suicidal ideation with low variability may be more likely to engage in suicidal behaviours with greater planning^{49,50}. As such, providing clinicians with an understanding of an individual's variability in suicidal ideation could help to tailor a more personalised suite of interventions.

The frequency of data collection is crucial in the context of MBC and associated tools where data is collected at varying time intervals, ranging from hours (e.g. ecological momentary assessments) or weekly/monthly (e.g. when prompted by a clinician at an appointment). Monitoring and data entry is driven by a multitude of factors including feasibility, ease of collection and motivation and the outcomes of interest (e.g. suicidal ideation, mood) which are important, dynamic and specific to that individual. However, research into the timeliness of data collection for an individual has been limited (see ref. 51 as an exception). As such, approaches that can provide timely prompts for information given an individual's specific circumstances would increase the level of personalised care.

Our analysis highlights that variability differs across the population and that predictions from a baseline observation are highly uncertain. The FUT values indicate how quickly information is lost and provide an approximate measure for how often an individual should be observed. These values tend to be lower when little information about a person is known (e.g. at baseline), or when someone has a highly variable course of illness. This implies that individuals should be observed more frequently during the early stages of care⁵² and when there is high observational variability. For example, the individual in Fig. 2d should be observed frequently (approximately weekly) compared to other individuals that had more stable trajectories.

Decisions regarding monitoring frequency need to consider several modelling and clinical factors not fully accounted for here. The recommendations assume a desired level of certainty (95% ETI range within 50% of the SIDAS range), which can be adjusted to increase or decrease the monitoring frequency. This implies that the FUT estimates should be considered as relative, rather than absolute. Furthermore, the emergence of suicidal ideation within individuals with no prior history of suicidal ideation necessitates ongoing observations. To improve our understanding of these phenomena, future work should focus on allowing for abrupt changes in an individual's state or incorporating dynamic parameters.

Our suggestion is to observe people when we know that our prior observations have limited predictive value, rather than to observe people at an absolute higher ideation level. This is a product of the random walk nature of the model parameterisation. If there was a dependency between current suicidal ideation scores and the future change in those scores, our recommendations may change.

A future goal for this work is to construct a modelling procedure that could be implemented using digital technologies within an MBC setting. We also want to expand this to several factors that are routinely monitored. The product would be a modelling approach that provides insights about the possible trajectories that an individual could take along with the summarisation of those trajectories—such as the individualised high-ideation probability, variability and recommended FUT—that could provide quick and informative insights for clinicians. We provide an illustrative example of the trajectory updating procedure in Fig. 3.

We acknowledge that there are several limitations in this work. Data are limited to those individuals who have engaged in clinical care using the Innowell platform. Furthermore, the data may not be missing at random as individuals were not obliged to enter their data into Innowell, which could lead to biases in the recorded level of suicidal thoughts.

The model formulation does not fully capture several aspects of suicidal ideation. The most important of which may be the binary aspect of suicidal ideation. Many individuals have no suicidal ideation throughout care, or they may transition from no suicidal ideation to a non-zero value. Similarly, other instantaneous movements or dynamic changes in their diffusion parameters are unlikely to be well-modelled. Model expansions could be made, including transitions to other qualitatively different states⁵³ or including time-varying parameters⁵⁴. We also do not explicitly include covariates at baseline within the model.

Model comparison via cross-validation is usually preferred to information criterion measures. However, a k -fold cross-validation procedure has a run time on the order of k . The sampling procedure typically takes several hours depending on the number of samples requested and the model parameterisation. Thus, testing many different model parameterisations would be infeasible. As such, we used an information criterion approach to allow us to explore models at a faster rate, at the expense of using a more accurate cross-validation technique.

Several simplifications are made to estimate the updated prediction trajectories in Fig. 3. In this case, we fixed the prior for the individual-level parameters using the maximum a posteriori estimates for the population-

level parameters. Thus, the population parameter uncertainties are not passed through. This is explained in Supplementary Methods 5.

Methods

Ethics

The Northern Sydney Local Health District Human Research Ethics Committees approved this study (HREC/17/HAWKE/480), and all participants gave online informed consent (via an opt-out process).

Participants

Participants were recruited from a group of individuals who presented for the first time to Headspace services ($N = 400$, 68.4%) in New South Wales, Queensland and South Australia between November 2018 and November 2022. We also include data from several other locations including Mind Plasticity ($N = 159$, 27.1%), Open Arms ($N = 13$, 2.2%), and the Butterfly Foundation ($N = 13$, 2.2%). All individuals presenting to these participating services were deemed eligible for this study, with the added requirement that they could access the Innowell platform. We included all individuals who used the Innowell platform with at least two completed observations for the suicidal ideation measure described below.

The Innowell platform

The Innowell Platform is an online platform that aims to assist in the assessment, management, and monitoring of mental ill health and maintenance of well-being^{3,55}. It is a web-based platform that can be accessed via traditional computing and mobile devices. The platform allows young people to complete web-based clinical assessments to understand their needs; explore their personalised dashboard of results; select from recommended care options (e.g. fact sheets, apps, e-tools and other web-based systems) to support their mental health and well-being; track their progress and share their dashboard and plan with clinicians to support care.

Measures

After being sent an invitation to join the Innowell platform, participants were asked to complete an initial assessment prior to their scheduled face-to-face appointment with a clinician. All assessment data are part of the functionality of the platform, which participants complete as part of standard clinical care through their service.

The initial clinical assessment covers a range of mental health concerns, as well as comorbid and associated risk factors. The assessment includes a range of biopsychosocial domains, such as mental health (i.e. psychological distress, depressed mood, anxiety, psychosis-like experiences, mania-like experiences and post-traumatic stress), suicidal thoughts and behaviours, social and occupational functioning, sleep-wake cycle, social connectedness, alcohol use, tobacco use, non-suicidal self-harm, physical health, eating behaviours and body image. It also assesses demographic information and a history of physical and mental health problems and treatment.

In addition to the initial questionnaire, individuals can complete a shorter follow-up assessment that included measures of suicidal ideation (SIDAS⁵⁶) and behaviours (C-SSRS⁵⁷), social functioning (SOFAS⁵⁸), overall mental health (Clinical Global Impressions Scale, CGI-S⁵⁹), overall health (single domain on EQ-5D-Y⁶⁰) and social support (Schuster's Social Support Scale, SSSS⁶¹). For more details about the measurements available within Innowell see Supplementary Methods 1.

We focus on the summed SIDAS score. SIDAS consists of five questions that focus on suicidal ideation. The questions ask individuals about: their frequency of suicidal thoughts, their ability to control those thoughts, whether they are tormented by those thoughts, and how much their suicidal thoughts impact their daily lives. They are also asked about their recent closeness to an attempt (0 = not close at all, 10 = made an attempt). All questions are asked on a 0–10 scale with higher values being more severe.

The recording of follow-up data was completed along with clinical care. The first summed SIDAS score was taken within the baseline questionnaire. All subsequent assessments are labelled with a time since baseline. As

individuals and clinicians were not required to enter data into Innowell the observations can be taken at various times throughout care.

Statistical analysis

The code for the model parameterisation and sampling scheme was written in Julia (version 1.8). All other analyses were written in R (version 4.3.1).

Our models were built using a data-driven model-building approach, where we started with a simple model parameterisation, and then increased complexity only when model performance warranted it. Such a model-building approach is useful to build simple models that can often have greater predictive performance. However, this model-building approach may leave out important features of the causal processes involved in suicidal ideation, and thus have limited ability in detailing our theoretical understanding of the underlying process.

We used continuous time models to deal with the irregularly time-spaced data. These models assume that the data follows a stochastic differential equation (SDE)^{39,40,62}. We explore two simple models known as a Wiener process and an Ornstein–Uhlenbeck process. The Wiener process is the simplest SDE, which corresponds to a continuous time random walk process, meaning that an individual's (m) trajectory will walk away from its current state in a random manner at a rate that is described by a diffusion parameter ($\phi_{m,2}$). The Ornstein–Uhlenbeck process is the continuous time equivalent of an autoregressive process, and increases the model complexity slightly by adding a deterministic term to the Wiener process that drives an individual's current state towards a long-term constant value ($\phi_{m,4}$) at a rate that is described by a drift parameter ($\phi_{m,3}$). All models have an initial baseline value ($\phi_{m,1}$) which is a parameterisation of the individual's initial suicidal ideation.

A novel addition from previous methods used on mental health data is that the modelled state values are converted from an unbounded continuous space to a discretised and bounded score using a logistic link function and a discretised normal distribution. This is implemented using a particle filter method, which allows for the fitting of complex models using simulations when analytical distributions are not available.

We explored several combinations of fixed and random effects for the model parameters. Where random effects are used, those parameters are assumed to follow a normal distribution^{42,63} that describes the population. In this way, the parameterisation allows for both population and individual-level parameters. More details of the model parameterisation can be found in Supplementary Methods 2.

The model was estimated within a Bayesian framework via posterior sampling⁶⁴. The aim of these techniques is to sample from the posterior distribution $p(\theta, |, y) \propto p(\theta)p(y|\theta)$, where θ is a set of parameters of interest that we would like to learn given some data y . The right-hand side informs us that our posterior knowledge is a function of our prior knowledge encoded in $p(\theta)$ which is updated with respect to the likelihood $p(y|\theta)$. We construct a scheme to sample from the posterior distribution that builds on the work by Wqvist et al.⁶⁵. This sampling scheme uses a correlated pseudo-marginal metropolis-hastings scheme to sample the individual parameters within a set of population parameters. The population parameters are then sampled using a semi-conjugate Gibbs sampling scheme. More details of this procedure are provided in Supplementary Methods 3.

Trajectories are a summary of the posterior distribution, where we use the median and equal-tailed credible intervals (ETI⁶⁶, a Bayesian equivalent to confidence intervals) as descriptors.

Predictions are constructed using the posterior predictive distribution. This involves simulating future trajectories many times while iterating over the parameter values drawn from the posterior distribution. More details on this procedure are provided in Supplementary Methods 4.

In all cases, we estimate the posterior predictive distributions for 60 days into the future from a reference observation. Sixty days was chosen as a reasonable period which was consistent with our typical time difference between observations, but it could be changed to account for other time periods of interest. We then summarise the distribution by calculating the median and the ETI on each day. We also provide two severity-related

statistics to summarise the posterior predictive trajectories. We calculate the probability that the individual will be in the high-ideation category (SIDAS > 20) on any day over the next 60 days (i.e. IHIP). Furthermore, we calculate the number of days that it would take for the 95% ETI to encompass at least 50% of the SIDAS range (i.e. 95% ETI range > 25) as an estimate for how long it takes to 'lose' information given the previous observations. We refer to this as the FUT throughout the paper. We also provide a measure of future observational variability (V), which is the expected ETI range 60 days from a baseline observation, along with the 80% highest density credible intervals (HDI) for that ETI range. We considered the 80% HDI range for V to be more parsimonious than wider measures as the distribution can be very wide-tailed and asymmetric. These summaries aim to provide a snapshot of the possible trajectories that an individual could take as well as their overall severity.

Model comparison

As a statistical model comparison tool, we compare the predictive ability of different models using the WAIC⁶⁷. The WAIC is similar to the Akaike information criteria as it estimates the out-of-sample predictive accuracy using a combination of the in-sample log-likelihood and a punishment for the degrees of freedom of the model^{46,47}. However, the WAIC is appropriate for Bayesian models where we have the posterior distribution of the log-likelihood rather than a maximum log-likelihood estimate. Furthermore, in hierarchical models, the degrees of freedom are typically less than the total number of parameters. The degrees of freedom for the WAIC are estimated as the variance of the posterior log-likelihood. WAIC is calculated on the deviance scale, and thus models with lower WAIC are preferred. While cross-validation methods provide a more robust measure of out-of-sample predictive accuracy, we considered such methods to be too computationally costly in our case.

Data availability

The datasets used and analysed during the current study are available from the corresponding author upon reasonable request.

Code availability

All relevant code for this paper can be found at [github.sydney.edu.au/mvar0005/dynamiclearning/](https://github.com/sydney.edu.au/mvar0005/dynamiclearning/).

Received: 24 October 2023; Accepted: 25 April 2024;

Published online: 07 June 2024

References

- Klein, A. et al. Remote digital psychiatry for mobile mental health assessment and therapy: MindLogger platform development study. *J. Med. Internet Res.* **23**, e22369 (2021).
- Torous, J. et al. Creating a digital health smartphone app and digital phenotyping platform for mental health and diverse healthcare needs: an interdisciplinary and collaborative approach. *J. Technol. Behav. Sci.* **4**, 73–85 (2019).
- Iorfino, F. et al. A digital platform designed for youth mental health services to deliver personalized and measurement-based care. *Front. Psychiatry* **10**, 1–9 (2019).
- Jellins, L. Assessment in the digital age: an overview of online tools and considerations for school psychologists and school counsellors. *J. Psychol. Couns. Sch.* **25**, 116–125 (2015).
- Chung, J. & Buchanan, B. A self-report survey: Australian Clinicians' attitudes towards progress monitoring measures. *Aust. Psychol. Soc.* **54**, 3–12 (2019).
- Kwan, B., Rickwood, D. J. & Telford, N. R. Development and validation of MyLifeTracker: a routine outcome measure for youth mental health. *Psychol. Res. Behav. Manag.* **11**, 67–77 (2018).
- Piwek, L., Ellis, D. A., Andrews, S. & Joinson, A. The rise of consumer health wearables: promises and barriers. *PLoS Med.* **13**, 1–9 (2016).
- Merikangas, K. R. et al. Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. *JAMA Psychiatry* **76**, 190–198 (2019).
- Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R. & Riemer, M. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. *Psychiatr. Serv.* **62**, 1423–1429 (2011).
- Scott, K. & Lewis, C. C. Using measurement-based care to enhance any treatment. *Cogn. Behav. Pract.* **22**, 49–59 (2015).
- Lambert, M. J. et al. Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clin. Psychol. Sci. Pract.* **10**, 288–301 (2003).
- Trivedi, M. H. & Daly, E. J. Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. *Drug Alcohol Depend.* **88**, S61–71 (2007).
- Harding, K. J. K., Rush, A. J., Arbuckle, M., Trivedi, M. H. & Pincus, H. A. Measurement-based care in psychiatric practice: a policy framework for implementation. *J. Clin. Psychiatry* **72**, 1136–1143 (2011).
- Hickie, I. B. et al. Right care, first time: a highly personalised and measurement-based care model to manage youth mental health. *Med. J. Aust.* **211**, S3–S46 (2019).
- Parikh, A., Fristad, M. A., Axelson, D. & Krishna, R. Evidence base for measurement-based care in child and adolescent psychiatry. *Child Adolesc. Psychiatry Clin. N. Am.* **29**, 587–599 (2020).
- Rognstad, K., Wentzel-Larsen, T., Neumer, S. P. & Kjøbli, J. A systematic review and meta-analysis of measurement feedback systems in treatment for common mental health disorders. *Adm. Policy Ment. Health.* **50**, 269–282 (2023).
- de Jong, K. et al. Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: a multilevel meta-analysis. *Clin. Psychol. Rev.* **85**, 102002 (2021).
- Iorfino, F. et al. Social and occupational outcomes for young people who attend early intervention mental health services: a longitudinal study. *Med. J. Aust.* **216**, 87–93 (2022).
- Hannan, C. et al. A lab test and algorithms for identifying clients at risk for treatment failure. *J. Clin. Psychol.* **61**, 155–163 (2005).
- Hatfield, D., McCullough, L., Frantz, S. H. B. & Krieger, K. Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clin. Psychol. Psychother.* **17**, 25–32 (2010).
- Walfish, S., McAlister, B., O'donnell, P. & Lambert, M. J. An investigation of self-assessment bias in mental health Providers. *Psychol. Rep.* **110**, 639–644 (2012).
- Graham, S. et al. Artificial Intelligence for mental health and mental illnesses: an overview. *Curr. Psychiatry Rep.* **21**, 116 (2019).
- Chekrou, A. M. et al. The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry* **20**, 154–170 (2021).
- Lutz, W., Schwartz, B. & Delgadillo, J. Measurement-based and data-informed psychological therapy. *Annu. Rev. Clin. Psychol.* **18**, 71–98 (2022).
- Lutz, W., Rubel, J. A., Schwartz, B., Schilling, V. & Deisenhofer, A. K. Towards integrating personalized feedback research into clinical practice: development of the trier treatment navigator (TTN). *Behav. Res. Ther.* **120**, 103438 (2019).
- Kwan, B. & Rickwood, D. J. A routine outcome measure for youth mental health: clinically interpreting MyLifeTracker. *Early Interv. Psychiatry* **15**, 807–817 (2021).
- Belsher, B. E. et al. Prediction models for suicide attempts and deaths: a systematic review and simulation. *JAMA Psychiatry* **76**, 642–651 (2019).
- Franklin, J. C. et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol. Bull.* **143**, 187–232 (2017).

29. McHugh, C. M. et al. Predictive modelling of deliberate self-harm and suicide attempts in young people accessing primary care: a machine learning analysis of a longitudinal study. *Soc. Psychiatry Psychiatr. Epidemiol.* **58**, 893–905 (2023).
30. Iorfino, F. et al. Predicting self-harm within six months after initial presentation to youth mental health services: a machine learning study. *PLoS One* **15**, 1–16 (2020).
31. Iorfino, F. et al. The temporal dependencies between social, emotional and physical health factors in young people receiving mental healthcare: a dynamic Bayesian network analysis. *Epidemiol. Psychiatr. Sci.* **32**, e56 (2023).
32. Skinner, A., Osgood, N. D., Occhipinti, J. A., Song, Y. J. C. & Hickie, I. B. Unemployment and underemployment are causes of suicide. *Sci. Adv.* **9**, eadg3758 (2023).
33. Nock, M. K. et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br. J. Psychiatry* **192**, 98–105 (2008).
34. Kleiman, E. M. et al. Digital phenotyping of suicidal thoughts. *Depress Anxiety* **35**, 601–608 (2018).
35. Rizk, M. M. et al. Variability in suicidal ideation is associated with affective instability in suicide attempters with borderline personality disorder. *Psychiatry* **82**, 173–178 (2019).
36. Sedano-Capdevila, A., Porras-Segovia, A., Bello, H. J., Baca-García, E. & Barrigon, M. L. Use of ecological momentary assessment to study suicidal thoughts and behavior: a systematic review. *Curr. Psychiatry Rep.* **23**, 41 (2021).
37. Cox, R. C., Brown, S. L., Chalmers, B. N. & Scott, L. N. Examining sleep disturbance components as near-term predictors of suicide ideation in daily life. *Psychiatry Res.* **326**, 115323 (2023).
38. Littlewood, D. L. et al. Short sleep duration and poor sleep quality predict next-day suicidal ideation: an ecological momentary assessment study. *Psychol. Med.* **49**, 403–411 (2019).
39. Oud, J. H. L. & Jansen, R. A. R. G. Continuous time state space modeling. *Psychometrika* **65**, 199–215 (2000).
40. Deboeck, P. R., & Preacher, K. J. No need to be discrete: a method for continuous time mediation analysis. *Struct. Equ. Modeling* **23**, 61–75 (2015).
41. de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L. & Hamaker, E. L. Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. *Front. Psychol.* **8**, 1–19 (2017).
42. Oravecz, Z., Tuerlinckx, F. & Vandekerckhove, J. A hierarchical Ornstein–Uhlenbeck model for continuous repeated measurement data. *Psychometrika* **74**, 396–418 (2009).
43. Driver, C. C. & Voelkle, M. C. Hierarchical Bayesian continuous time dynamic modelling. *Psychol. Method* **23**, 774–799 (2018).
44. Merton, R. C. & Samuelson, P. A. (eds) *Continuous-Time Finance: Revised Edition* (Basil Blackwell, 1992).
45. Watanabe, S. A widely applicable bayesian information criterion. *J. Mach. Learn. Res.* **14**, 867–897 (2013).
46. Gelman, A., Hwang, J. & Vehtari, A. Understanding predictive information criteria for Bayesian models. *Stat. Comput.* **24**, 997–1016 (2014).
47. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2017).
48. Bermanke, J. A., Stanley, B. H. & Oquendo, M. A. Toward fine-grained phenotyping of suicidal behavior: the role of suicidal subtypes. *Physiol. Behav.* **176**, 139–148 (2018).
49. Nock, M. K. et al. Risk factors for the transition from suicide ideation to suicide attempt: results from the army study to assess risk and resilience in servicemembers (Army STARRS). *J. Abnorm Psychol.* **127**, 139–149 (2019).
50. Miranda, R., Ortin, A., Scott, M. & Shaffer, D. Characteristics of suicidal ideation that predict the transition to future suicide attempts in adolescents. *J. Child Psychol. Psychiatry* **55**, 1288–1296 (2014).
51. Adolf, J. K., Loossens, T., Tuerlinckx, F. & Ceulemans, E. Optimal sampling rates for reliable continuous-time first-order autoregressive and vector autoregressive modeling. *Psychol. Methods* **26**, 701–718 (2021).
52. Capon, W. et al. Clinical staging and the differential risks for clinical and functional outcomes in young people presenting for youth mental health care. *BMC Med.* **20**, 1–10 (2022).
53. Hamaker, E. L., Grasman, R. P. P. & Kamphuis, J. H. Modeling BAS dysregulation in bipolar disorder: illustrating the potential of time series analysis. *Assessment* **23**, 436–446 (2016).
54. Ryan, O., Haslbeck, J. & Waldorp, L. Non-stationarity in time-series analysis: modeling stochastic and deterministic trends. <https://doi.org/10.31234/osf.io/z7ja2> (2023).
55. Capon, W. et al. Characterising variability in youth mental health service populations: a detailed and scalable approach using digital technology. *Australas. Psychiatry* **31**, 295–301 (2023).
56. Van Spijker, B. A. J. et al. The suicidal ideation attributes scale (SIDAS): community-based validation study of a new scale for the measurement of suicidal ideation. *Suicide Life Threat. Behav.* **44**, 408–419 (2014).
57. Posner, K., Brown, G. K. & Stanley, B. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am. J. Psychiatry* **168**, 1267–1277 (2011).
58. Goldman, H. H., Skodol, A. E. & Lave, T. R. Revising axis V for DSM-IV: a review of measures of social functioning. *Am. J. Psychiatry* **149**, 1148–1156 (1992).
59. Busner, J. & Targum, S. D. Global impressions scale: applying a research tool in practice. *Psychiatry (Edgmont)* **4**, 28–37 (2007).
60. Wille, N. et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual. Life Res.* **19**, 875–886 (2010).
61. Schuster, T. L., Kessler, R. C. & Aseltine, R. H. Supportive interactions, negative interactions, and depressed mood. *Am. J. Community Psychol.* **18**, 423–438 (1990).
62. Arnold, L. *Stochastic Differential Equations: Theory and Applications*. SIAM Review (Wiley-Interscience, 1974). <https://doi.org/10.1137/1018036>.
63. Driver, C. C., Oud, J. H. L. & Voelkle, M. C. Continuous time structural equation modeling with r package ctsem. *J. Stat. Softw.* **77**, 1–35 (2017).
64. Gelman, A. et al. *Bayesian Data Analysis* (Chapman and Hall/CRC, 1995). <https://doi.org/10.1201/9780429258411>.
65. Wikqvist, S., Golightly, A., McLean, A. T. & Picchini, U. Efficient inference for stochastic differential equation mixed-effects models using correlated particle pseudo-marginal algorithms. *Comput. Stat. Data Anal.* **157**, 107151 (2021).
66. Kruschke, J. K. *Doing Bayesian Data Analysis: A tutorial with R, JAGS, and Stan, 2nd Edition*. (Elsevier Inc., 2015).
67. Watanabe, S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* **11**, 3571–3594 (2010).

Acknowledgements

This work was supported by the Medical Research Future Fund Applied Artificial Intelligence Research in Health Grant (MRFAI000097, ‘Explainable machine learning to improve youth mental health care’), and NHMRC Australia Fellowship (511921 awarded to I.B.H.). M.V. was supported by philanthropic funding from The Johnston Fellowship and from other donor(s) who are families affected by mental illness and wish to remain anonymous. F.I. was supported by an NHMRC EL1 Investigator Grant (GNT2018157). I.B.H. is supported by a NHMRC L3 Investigator Grant (GNT2016346). The funders played no role in the study design, data collection, analysis and interpretation of data or the writing of this manuscript.

Author contributions

M.V. was responsible for writing the manuscript and code. M.V., F.I., I.B.H., S.C., R.M., R.O. and A.P. provided substantial input into initial scoping and further iteration of the models such that they would both be statistically

rigorous and address the needs of clinicians. E.S. participated in the collection of data and provided feedback on the results. J.S. and A.S. provided important feedback on the manuscript that provided substantial improvement.

Competing interests

I.B.H. is the Co-Director, Health and Policy at the Brain and Mind Centre (BMC) The University of Sydney. The BMC operates an early-intervention youth service at Camperdown under contract to Headspace. He is the Chief Scientific Advisor to, and a 3.2% equity shareholder in, InnoWell Pty Ltd, which supports the transformation of mental health services internationally through the use of innovative technologies. E.S. is a Principal Research Fellow at the Brain and Mind Centre The University of Sydney. She is a Discipline Leader of Adult Mental Health, the School of Medicine, University of Notre Dame, and a Consultant Psychiatrist. She was the Medical Director, of the Young Adult Mental Health Unit, St Vincent's Hospital Darlinghurst until January 2021. She has received honoraria for educational seminars related to the clinical management of depressive disorders supported by Servier, Janssen and Eli-Lilly Pharmaceuticals. She has participated in a national advisory board for the antidepressant compound Pristiq, manufactured by Pfizer. She was the National Coordinator of an antidepressant trial sponsored by Servier.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44184-024-00071-0>.

Correspondence and requests for materials should be addressed to Mathew Varidel.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024