

Supplementary Online Content

Chien I, Enrique A, Palacios J, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. *JAMA Netw Open*. 2020;3(7):e2010791.
doi:10.1001/jamanetworkopen.2020.10791

eMethods. Detailed Methods

eTable 1. List of Tools Available on the SilverCloud Health iCBT Platform

eTable 2. Description of Sections Users Engaged With Over a 14 Week Period on the SilverCloud Platform

eTable 3. Estimated Class-Specific Change in PHQ-9 Over Time

eReferences

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Detailed Methods

Intervention

“*Space from depression and anxiety*” is a programme aimed at addressing symptoms of anxiety and depression, given the high comorbidity between both presentations. The content and the structure of the programme is based on evidence-based CBT principles and it follows the NICE guidelines for the treatment of these conditions (NICE, 2006). It is composed of eight modules that comprise the core depression modules (i.e. psychoeducation, cognitive restructuring, behavioral activation) along with a module about worry management. Programme content can be accessed in a non-linear fashion, although a linear path is recommended. Patients work through the program content at their own pace, and in their own time, with the recommendation to complete one module each week within over an eight-week period. Each module incorporates different types of content, such as informational text, quizzes, interactive activities, videos and others, that can be observed in Table 1. Furthermore, the supporters can also unlock specific modules of content to cover specific needs patients may have, such as sleep hygiene, stress management or self-esteem.

The programmes are available in 'self-help' mode or 'supported' mode where the patient is assigned a therapist who acts as their supporter. The supporter monitors the patient's activity on a regular basis (i.e. number of logins, page views, tools used, clinical scores), providing asynchronous support through messages on the platform. The role of the supporter is aimed at providing guidance and giving feedback on completed exercises. By default, the content of the activities completed by patients is shared with their supporters, which allows supporters to view users' goals for the week, key messages and progress points. If patients wish to share more content with their supporter, they can share their journal entries as well. The platform provides the facility for the patient to complete clinically validated outcome measures where appropriate.

Data description and preprocessing. Data was taken from adult patients enrolled on the SilverCloud platform from 2015-2019. SilverCloud Health is a leading provider of internet-based psychological interventions for different mental health conditions, such as depression, anxiety and stress. The core programs of the SilverCloud suite are: Space from Depression, Space from Anxiety, and Space from Stress. The programs are developed under the principles of cognitive behavioral therapy and follow the NICE guidelines for the treatment of each specific condition (Richards et al., 2014; Richards et al., 2018). These internet-based Cognitive behavioral therapy (iCBT) programs are delivered on a Web 2.0 platform using media-rich interactive content, in order to improve the user experience.

The database consisted of data from 239,043 patients. Of these patients, we restricted analysis to 67,468 patients who were assigned to the “Space from Depression and Anxiety” program. Of these, we removed 12,865 who had no supporter assigned to them and restricted the analysis to 54604 patients using the platform.

The database contains the following:

- Audit trail information on user interactions (e.g. page-views, login events, activity updates).
- User Data from completed exercises (e.g. journal entries, mood monitor entries).
- Messages between patient and therapist.
- Data from clinical outcome measures.

No personally identifying user information was included. e.g. name, email address and date of birth.

Prior to analysis, the following steps were taken by SilverCloud to anonymize the data:

- All web URLs were removed if associated with an individual patient.
- Text entries (e.g. journal entries, messages between patient and therapist) are not included. These were replaced with a field indicating character length of the original text.
- Specific dates relating to user data were replaced with an offset in seconds from the date they joined the platform. The year a user joined the platform was retained.
- All database IDs were mapped to different numbers using a private, non-sequential hashing function.
- Specific Information about a patient's referral pathway to SilverCloud was not included (e.g. the name of the University or health service). The country and sector of origin was retained.

The data is transferred as a secured SQLite database file. This is stored in a secure data center. To enable the machine learning algorithms we extract the data from the tables in the database into a series of CSV files using SQL and vectorized pandas functions. The majority of the data is extracted from one table that holds the 'analytic events' (i.e. the events captured through platform use). Additional data for more specific requirements (like the various assessment scores) is gleaned from other normalized tables. To facilitate understanding a certain amount of de-normalization occurs during this extraction process, though we also derive minimal normalized CSV files to reduce the memory footprint in the machine learning algorithms where necessary.

For our class subtyping models, we used the analytic events data, counting any such item as an interaction with the platform. We filtered the data to only the first 14 weeks of therapy for each user. Our class subtypes are therefore based purely in patient engagement with the platform. We analyzed these subtypes with data from clinical outcome measures, which indicate a patient's mental health progression.

Model experimentation. We explored two model formulations to infer the presence and number of (latent) subtypes of user engagement.

1. A Latent Variable Mixed Model where we assume independence conditional on membership to some discrete latent class
2. A Hidden Markov model where we assume serial conditional dependence structure on a sequence of engagement observations over 14 weeks

These are unsupervised learning tasks where we want to infer hidden structure from patterns of engagement over time.

We specify the following variables and parameters which apply to both the latent variable mixture models and the hidden Markov models:

- i , representing the patient/subject
- t , the time in weeks
- $T = 14$, the total number of weeks
- k , classes (subtypes) that a patient/subject may belong to
- K , the total number of classes (which is variable based on experimentation)
- c_i denoting the class membership of patient/subject i , $c_i \in \{1, \dots, K\}$

Latent variable mixture models. We explored two variations of latent variable mixture model for class subtyping:

- i. A latent class model based on binary engagement within the first 14 weeks of therapy
- ii. A latent class model based on engagement by section within the first 14 weeks of therapy

In the latent class model, the patients are assumed to belong to one of K discrete where the number of classes, size of classes and class membership is unknown. The classes are categories of a categorical latent variable. Latent class models are used when dichotomous or polytomous observations y are observed on each patient i .

For a set of i patients at time t , the likelihood of each observation y_{it} is:

$$p(y_{it} = 1 | c_i = k) = \frac{\exp(\alpha_k + \beta_k t)}{1 + \exp(\alpha_k + \beta_k t)}$$

The observation, y , is assumed to be conditionally independent given membership to a given latent class.

α and β are weights for the intercept and time respectively. This model formulation can be written in terms of a longitudinal logistic regression model in the form:

$$\text{logit}\{p(y_{it} = 1 | c_i = k)\} = \alpha_k + \beta_k + \xi_i$$

where β_k represents time-dependent class effects, where α_k represents the fixed class effects, and ξ_i represents patient-specific random effects.

The latent variable π is a K -dimensional prior probability vector which mixes individual Gaussian distributions, each characterized by a mean μ_k and variance σ_k^2 . It represents the prior probability that a patient i belongs to a class k .

We define the prior on $\pi \in [0,1]$ such that $\sum_{k=1}^K \pi_k = 1$ to be

$$p(\pi) = \text{Dirichlet}\left(\frac{1}{K}, \dots, \frac{1}{K}\right)$$

We define the prior on μ_k to be

$$p(\mu_k) = \text{Normal}(\mu_k | 0, I)$$

We define the prior on σ_k^2 to be

$$p(\sigma_k^2) = \text{InverseGamma}(\sigma_k^2 | a, b)$$

The posterior distribution of the latent variables is:

$$E_z[\ln p(y, c | \mu, \pi)] = \sum_{i=1}^N \sum_{k=1}^K E[c_i = k] \{ \ln \pi_k + \sum_{t=1}^T [y_{it} \ln \mu_{ik} + (1 - y_{it}) \ln (1 - \mu_{ik})] \}$$

where $E[c_i = k]$ is the posterior probability of component k given datapoint y_{it} .

We evaluate these probabilities in the inference step, by adjusting parameters to achieve maximum likelihood estimations for patient class membership. These models are implemented in Python, using the package Pyro.

Hidden Markov models. A Hidden Markov Model (HMM) is an extension of the latent class models described above. For an HMM, we explicitly condition the latent state x_t (latent state at time t) to be conditional on the x_{t-1} . We express the transitions between states with a transition matrix, which contains probabilities of moving from one hidden state to another between time $t - 1$ and time t . We explored three variations of HMM for class subtyping:

- i. An HMM based on binary engagement within the first 14 weeks of therapy.
- ii. An HMM based on engagement by sections within the first 14 weeks of therapy, transformed into a single observation by week per user.
- iii. An HMM based on engagement by sections within the first 14 weeks of therapy, treated as multivariate observations.

This involves a slightly different model formulation and implementation of the HMM.

We specify the following variables and parameters, which apply for all three variations of HMM unless otherwise specified:

- For model variation (i), binary engagement of users on the platform was used as observed input data. Observation was defined as 1 if the user had interacted (engaged) with the SilverCloud platform at all in the specified time point (by week) and 0 if not.
 $y = (y_{i1}, y_{i2}, \dots, y_{iT})$, an observation sequence with observations $\in \{0, 1\}$ for each subject i , for times $t \in \{1, T\}$

For model variation (ii), engagement for a section was defined as if the user had interacted with that section of the platform in that week and 0 if not. These binary values were concatenated into a binary string, which was then translated into a single, base-10 integer. For example, if using observations from sections 1, 2, and 3 and the user had only engaged with section 2, the string would be '010' and the resulting integer observation would be 2. These observations capture the distinctions between users who engage with certain sections. This variation uses the same model formulation as detailed above, with the only difference being the input data.

$y = (y_{i1}, y_{i2}, \dots, y_{iT})$, an observation sequence with observations as described for each subject i , for times $t \in \{1, T\}$

For model variation (iii), engagement for sections by users on the platform was treated as multivariate observations, with each section occupying a different dimension of each observation. Engagement was 1 if the user engaged with a section in that week, 0 if not. This variation requires a modified model formulation to account for the multivariate observations.

$y = (\{y_{i1}^{(s)}\}, \dots, \{y_{iT}^{(s)}\})$ an observation sequence with observations $\in \{0, 1\}$ for each subject i , at each dimension s , for times $t \in \{1, T\}$.

- $x = (x_{i1}, x_{i2}, \dots, x_{iT})$, a latent state sequence with states $b \in \{1, 2, \dots, B\}$ for each subject i , for times $t \in \{1, T\}$

To accommodate our class subtyping, the number of latent states, B , is equivalent to the number of intended hidden states (H) multiplied by number of classes (K).

- $\pi = \{\pi_b\}$, a B -dimensional initial probability vector, representing the probability of starting from latent state b : $\pi_b = P(x_{i1} = b)$; $b \in \{1, \dots, B\}$

We constructed a specific initialization of the initial probability vector to perform class subtyping. The initial probability of a state can only be nonzero for states considered the initial states, which would represent the starting latent state of an observation within a class. The starting probability is Dirichlet over all the initial states and set to 0 for all other states.

- $A = \{a_{br}\}$, a transition matrix of size $B \times B$ where a_{br} is the probability of moving from latent state b at time $t-1$ to latent state r at time t :

$$a_{br} = P(x_{it} = r | x_{it-1} = b); b, r \in \{1, \dots, B\}$$

Structure was also imposed on the transition matrix to define classes; transitions can only occur between states within a class, not from a state of one class to a state of another.

Therefore, those transitions were initialized with a Dirichlet probability distribution, over the states within a class, while all others were assigned a zero value.

- $Q = \{q_b(m)\}$, an emission matrix of size $B \times M$ where $q_b(m)$ is the probability of the latent state b emitting the observation m :

$$q_b(m) = P(y_{it} = m | x_{it} = b); b \in \{1, \dots, B\}; m \in \{1, \dots, M\}$$

The emission matrix is randomly initialized.

Model parameter learning. All formulations of the Hidden Markov model were implemented in Python 3.6, using the package `hmmlearn`. Input data was cleaned and anonymized, and transformed into usable matrices. We modified the package `hmmlearn` to support multivariate input data with multinomial (discrete) emissions. HMM parameters were initialized as detailed above. HMMs were fit using the Baum-Welch expectation maximization algorithm and decoded (state assignment) using the Viterbi algorithm [1]. Small modifications were made to the update of the emission matrix to accommodate our multivariate sections observations, for model variation (iii).

The Markov assumption states that the hidden state transition probability at time t only depends on the hidden state at the previous time point $t - 1$:

$$P(x_{it} | x_{it-1}, \dots, x_{i1}) = P(x_{it} | x_{it-1})$$

The observation at time t is only dependent on the current hidden state, not the previous hidden states or observations:

$$P(y_{it} | y_{it-1}, \dots, y_{i1}, x_{it}, \dots, x_{i1}) = P(y_{it} | x_{it})$$

In the case where we have multiple sections, we assume that there are S independent parallel sequences. We have one transition matrix A , but several emission matrices Q_1, \dots, Q_S that are updated in parallel. s is the dimension of the observation, S is the total number of dimensions (in our case, the number of sections we are including in our analysis). We assume that the observations along different dimensions (sections) at a given time point t are independent of each other given the latent state at t . Therefore, the probability of observing y_{it} at state x_{it} is the product of the probabilities of observing each dimension of the observation y_{it} is:

$$P(y_{it} | x_{it}) = q_{x_{it}}(y_{it}) = \prod_{s=1}^S P(y_{it}^{(s)} | x_{it}) = \prod_{s=1}^S q_{x_{it}}^{(s)}(y_{it}^{(s)})$$

The likelihood of the observation sequence of patient i given the model parameters $\Lambda = \{\pi, A, Q_1, \dots, Q_S\}$ for the HMM is:

$$\begin{aligned} P(Y_i | \Lambda) &= \sum_x P(Y_i | x, \Lambda) P(x | \Lambda) \\ &= \sum_x P(x_{i1} | \Lambda) P(y_{i1} | x_{i1}, \Lambda) \prod_{t=2}^T [P(x_{it} | x_{it-1}, \Lambda) P(y_{it} | x_{it}, \Lambda)] \end{aligned}$$

When we use multiple sections in our observations, the likelihood is:

$$P(Y_i | \Lambda) = \sum_x P(x_{i1} | \Lambda) \prod_{s=1}^S P(y_{i1}^{(s)} | x_{i1}, \Lambda) \prod_{t=2}^T \left[P(x_{it} | x_{it-1}, \Lambda) \prod_{s=1}^S P(y_{it}^{(s)} | x_{it}, \Lambda) \right]$$

$$\begin{aligned}
&= \sum_x P(x_{i1}|\Lambda) \prod_{s=1}^S P(y_{i1}^{(s)} | x_{i1}, \Lambda) \prod_{t=2}^T \left[P(x_{it}|x_{it-1}, \Lambda) \prod_{s=1}^S P(y_{it}^{(s)} | x_{it}, \Lambda) \right] \\
&= \sum_x \pi_{x_{i1}} \prod_{s=1}^S q_{x_{i1}}^{(s)}(y_{i1}^{(s)}) \prod_{t=2}^T \left[a_{x_{it-1}x_{it}} \prod_{s=1}^S q_{x_{it}}^{(s)}(y_{it}^{(s)}) \right]
\end{aligned}$$

The summation over x is over all possible states x . The subclasses K are codified in the initializations of the parameters. The calculation of this likelihood is computationally intensive, so we use the Baum-Welch expectation maximization algorithm to find the maximum likelihood estimate of the parameters for a HMM, given a set of observations. We aim to optimize π , the initial probability vector; A , the transition matrix; and Q , the emission matrix. We made slight adjustments to the forward and backward procedures of the Baum-Welch algorithm with multiple dimensions in mind. For the following equations we are dropping the patient indicator variable i for simplicity.

Baum-Welch makes use of a forward-backward procedure. The forward procedure, which is the probability of seeing observations y_1, \dots, y_t and being in state b at time t is defined as:

$$\alpha_b(t) = P(y_1, \dots, y_t, x_t = b | \Lambda)$$

This is calculated recursively with:

$$\begin{aligned}
\alpha_b(1) &= \pi_b q_b(y_1) \\
\alpha_b(t+1) &= q_b(y_{t+1}) \sum_{r=1}^B \alpha_r(t) a_{rb}
\end{aligned}$$

The backward procedure, which is the probability of seeing observations y_t, \dots, y_T given the state b at time t is defined as:

$$\beta_b(t) = P(y_{t+1}, \dots, y_T | x_t = b, \Lambda)$$

This is calculated recursively with:

$$\begin{aligned}
\beta_b(T) &= 1 \\
\beta_b(t) &= \sum_{r=1}^B \beta_r(t+1) a_{br} q_r(y_{t+1})
\end{aligned}$$

To perform inference, we use temporary update variables:

$$\begin{aligned}
\gamma_b(t) &= P(x_t = b | Y, \lambda) = \frac{P(x_t = b, Y | \lambda)}{P(Y | \lambda)} = \frac{\alpha_b(t) \beta_b(t)}{\sum_{r=1}^B \alpha_r(t) \beta_r(t)} \\
\xi_{br}(t) &= P(x_t = b, x_{t+1} = r | Y, \lambda) = \frac{P(x_t = b, x_{t+1} = r, Y | \lambda)}{P(Y | \lambda)} \\
&= \frac{\alpha_b(t) a_{br} \beta_r(t+1) q_r(y_{t+1})}{\sum_{b=1}^B \sum_{r=1}^B \alpha_b(t) a_{br} \beta_r(t+1) q_r(y_{t+1})}
\end{aligned}$$

Finally, the parameters of the model are updated as the following:

$$\begin{aligned}
\pi_b^* &= \gamma_b(1) \\
a_{br}^* &= \frac{\sum_{t=1}^T \xi_{br}(t)}{\sum_{t=1}^T \gamma_b(t)} \\
q_b^*(v_d) &= \frac{\sum_{t=1}^T \mathbf{1}_{(y_t = v_d)} \gamma_b(t)}{\sum_{t=1}^T \gamma_b(t)}
\end{aligned}$$

For model inference on multivariate data (when we incorporate the separate sections data), the calculation of the emission probability $q_b(y_t)$ is handled as defined above, with all else remaining the same. The exception is in the update of the emissions matrix; each emission matrix (for each dimension of the multivariate observation) is updated independently. Updates occur for all D possible observation values $v_d^{(s)}$ with d indexing the value, for each section s , while in state b :

$$\begin{aligned}
q_b^{*(s)}(v_d^{(s)}) &= \frac{\sum_{t=1}^T \mathbf{1}_{(y_t^{(s)} = v_d^{(s)})} \gamma_b(t)}{\sum_{t=1}^T \gamma_b(t)} \\
\mathbf{1}_{(y_t^{(s)} = v_d^{(s)})} &= \begin{cases} 1 & \text{if } y_t^{(s)} = v_d^{(s)} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

The class subtype for each user is determined by the latent state of the first timepoint observation.

We experimented with different numbers of hidden states and numbers of classes, and determined the most suitable classification of patients using the Bayesian Information Criterion (BIC) [2] of the model, as well as qualitative analysis. HMMs were trained and evaluated on the full dataset.

The BIC is used to compare models with different number of classes. It is a penalized log likelihood, which penalizes the model for size and complexity:

$$BIC = -2 \log L + p \log(T)$$

Where T = the length of the time-series (254743 users x 14 time points)

p = number of independent parameters in the model. P can be determined by

$$p = m^2 + km - 1$$

where m = number of states in the HMM and k = number of parameters of the underlying distribution of the data (2 for a binomial distribution).

To investigate differences in subtypes in terms of clinical outcomes, we used a longitudinal linear regression model where subtype assignment, week and an interaction between subtype assignment and week were used as predictors of PHQ-9 or GAD-7 at a given time point. These results are presented in Table 2. Furthermore, to test for consistency and the assumption of missingness at random, we also ran this analysis restricting it to users who had at least 3 PHQ-9 assessments completed [n=31,466]. We found consistent estimates in terms of the estimated mean improvement in depression and anxiety over time within each subgroup (Table A2)

We also investigated whether there were particular sections on the platform that certain subtypes used in the first 2 weeks on the platform. These findings are summarised in Table 1. To achieve this, we used a logistic regression model where we defined a binary variable for a given class (value=1) vs. all other classes (value=0).

eTable 1: List of tools available on the SilverCloud Health iCBT platform			
My TFB Cycles	Understanding My Problems	Thoughts & Worries Quiz	Boosting Behavior
Understanding My Situation	Take Control	How Do You Manage Your Thoughts	Impulse Spending Quiz
Mood Monitor	Work Satisfaction	Staying Well Plan	My Hope Coping Statements
Goals	Grief Myths Facts Quiz	Time Management Tool	Pros & Cons of Uncertainty
Staying in the Present	Sleep Diary	Lifestyle Choices Chart	TRAC
Activity Scheduling	Wordle	My Early Warning Signs	Having A Healthy Lifestyle
Emotions and your Body	Eating and Emotions	Applying new skills online quiz	Myths and Facts Quiz
Worry Tree	You and your anxiety	Me and my thoughts	My Triggers for Impulse Spending
Activities List	My Symptoms Of Grief	Theory A / Theory B	Facing Your Financial Fears Quiz
Hierarchy of Fears	Chronic Pain Myths and Facts	Lifestyle Tips - Exercise	Stress Response
Anxiety Myths and Facts	Coping Mechanisms	Tips for Wellbeing	Worry: Challenging Beliefs
What's your lens? What's your thinkin..	My Grief Emotions	Lifestyle Tips - Diet	Do I Have a Problem? Quiz
My Worries	Body Image Tips	Lifestyle Tips - Sleep	Anxiety & Worry Quiz
Challenge Thoughts	Me and My Thoughts	Vicious Flower	Safety Behaviours
Staying Well	Attitudes to Self Questionnaire	Getting Motivated	At-Risk Assessment
Helpful Thoughts	Resilience Star	Mood and Behaviour	Linking Worry and Anxiety Quiz
Getting Started Quiz	Anxiety Myths and Facts Quiz	Behaviour Experiments	My Triggers
Boosting Behaviour	My Past Losses	Perinatal Myths & Facts Quiz	Staying Well Financially
Your Body Your Mood	Stress Symptoms	Balanced Optimism	Behavioural Activation Quiz
Anxious Thoughts & Worry Quiz	Restoring Myself		Pausing Big Ideas

eTable 1: List of tools available on the SilverCloud Health iCBT platform (continued)

Depression: Myths and Facts	Confronting My Loss	Setbacks	Checklist
Your Life Areas	Doodle Goals	OCD Basics Quiz	Understanding My Anxiety
My Strengths	What Do You Know About Stress?	Diabetes Activities	Understanding Uncertainty
Graded Exposure Quiz	ORS	CVD Quiz	My Exercising Habits
Life Events Quiz	My Values	Crisis Plan	Am I At-Risk? Quiz
Stress - An Overview Quiz	Doodle Draw	Medications	Behavioural Experiments
Your Self Quiz	My Decisions	Care Plan	Do your behaviors match your values?
Life Style Choices Chart	My Sessions	Activities for those with Bipolar	Embracing Uncertainty
Coping Strategies	My Life Roles	What is the cost of OCD?	Enjoyable Activities
Problem Solving	Feelings Explorer	Overcoming Guilt and Shame	Responsibility Pie
Life Areas	What's Your Thinking Style?	Sleep Measure	Being a Support Person Quiz
Mindfulness	Diabetes Myths & Facts Quiz	Activities	Affirmations
Core Beliefs	txtspeak	OCD Development Quiz	Facing Your Fears Quiz
Balance in Self-Care	Getting to Know You	Problem Solving for Staying Safe	What Are Your Triggers
Coping Style Quiz	Mental Health Quiz	Positive Mental Health Quiz	Your Thinking Style
Sleeping Tips	Website Reviews	Learning from Experience	

eTable 2: Description of sections users engaged with over a 14 week period on the SilverCloud Platform.

For the analysis we excluded the sections 'Home', 'Downloads' as these had little variation between users and 'Resources' and 'Therapist' as these are therapist-specific sections

Section	Description
Home	All Homepage related activity. Users can perform some limited tasks here. Message supporter, Navigate to their program etc.
Tools	Tools: interactive activities where users are asked to apply specific techniques or information learnt through the intervention and take learnings from content and other parts of the platform (i.e. quizzes, mindfulness audios, mood tracker)
Content	Reading material in the program which mainly requires viewing material on the platform
Journal	A diary where the user can fill out free text
Todo'	Where people can fill out a 'to-do' list at the end of every module,
Likes	Every content page has a 'like' button
Profile	Where a user can fill out name, birthday and other information
Progress Point	A section on the last page of every module where people are asked to reflect on what they have learned. They are asked how much they agree/disagree with a series of questions 1. I found this module interesting. 2. This module was relevant to me. 3. This module was helpful to me in some way. 4. The program is supporting me to make progress towards my goals. 5. Any other feedback about this module? (Optional)
Review	The review page where they can communicate with their human supporter, review a summary of their SilverCloud activity, and share/unshare their activity with their supporter
Take-home point	A Section on the module summary page where clients can note down 'key messages' from the module
Therapist	Therapist related activity events. Not applicable
Resource	Not applicable
Measures	Filling out and viewing clinical measures of PHQ-9 and GAD-7
Downloads	Downloading any content/pdfs etc.

eTable 3: Estimated class-specific change in PHQ-9 over time. Analysis was restricted to patients who had at least 3 PHQ-9 measures.

Numbers in brackets represent 95% confidence intervals

Engagement Subtype	Estimated Mean Initial PHQ-9			Estimated Change in PHQ-9 after 14 weeks			Cohen d's Effect Size for Change in PHQ-9	
	Mean Score	LCL	UCL	Mean 14-week change	LCL	UCL	Cohen's d	SE
Baseline: Class 1	13.55	13.38	13.73	-5.11	-4.79	-5.26	(baseline)	
Class 2	12.41	12.03	12.80	-4.41	-3.87	-4.92	-0.46	0.014
Class 3	12.72	12.34	13.11	-6.64	-6.13	-7.15	-0.46	0.014
Class 4	12.99	12.55	13.34	-5.44	-4.71	-5.94	-0.61	0.021
Class 5	12.52	12.11	12.92	-5.88	-4.84	-6.38	-0.73	0.018

Code Availability

Code is available on request from the corresponding author

eReferences

1. Bilmes JA. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute*. 1998;4(510):126.
2. Schwarz G. Estimating the dimension of a model. *The annals of statistics*. 1978;6(2):461-4.