

ORIGINAL ARTICLE

High-frequency measurement of depressive severity in a patient treated for severe treatment-resistant depression with deep-brain stimulation

S Sani¹, J Busnello², R Kochanski¹, Y Cohen³ and RD Gibbons⁴

Although there have been previous studies of deep-brain stimulation (DBS), we present, to our knowledge, the first example of high-frequency depressive severity measurement-based DBS treatment in particular and psychiatric treatment in general. Daily post-surgical e-mail prompts for a period of 6 months resulted in 93 administrations of a computerized adaptive test (CAT) of depression severity (CAT-Depression Inventory or CAT-DI) via the internet. There was an average of 3.37 weekly measurements with an average separation of 2.12 days. No additional incentive was provided to the patient for completing the adaptive tests. The patient is a 55-year-old female with six psychiatric hospitalizations for depression, two suicide attempts, marginal response to eight electroconvulsive therapy (ECT) treatments and 35 psychotropic medications. We report results after high-frequency stimulation of the superolateral branch of the medial forebrain bundle. The CAT-DI was used for daily assessments before, during and after (remotely in response to an e-mail prompt) the DBS procedure. Two follow-up Hamilton Depression Scales (HAM-Ds) were also collected. Response to treatment varied markedly, with a decrease from severe (> 75) to mild (60), which is three times the size of the uncertainty level. Although the HAM-D scores decreased, they missed the more complete temporal pattern identified by CAT-DI daily monitoring. We demonstrated feasibility of daily depressive severity measurement at high levels of precision and compliance. Clinician ratings confirm the general pattern of treatment benefit, but mask the marked variability in mood and more marked periods of benefit and decline.

Translational Psychiatry (2017) **7**, e1207; doi:10.1038/tp.2017.145; published online 15 August 2017

INTRODUCTION

Although there have been previous studies of deep-brain stimulation (DBS) for depression,¹ there is a general lack of consensus on the optimal way to measure outcome. To our knowledge, the presented case study is the first example of successful high-frequency measurement-based DBS treatment for a patient with treatment-resistant depression (TRD). Measurements of depressive severity were taken before surgery, during surgery and daily during a 6-month follow-up period. To eliminate response bias associated with the repeated administration of the same questions on a traditional fixed-length test, we used a novel computerized adaptive test (CAT) based on multidimensional item response theory (IRT) for depression,² which allows the questions to adapt to the patient's changing level of depressive severity. This paradigm has been previously described³ and asks different questions on each repeat administration. To our knowledge, this is the first example of high precision, high-frequency adaptive measurement-based treatment in psychiatry, which is the primary rationale for this study.

Deep-brain stimulation

DBS involves electrical stimulation of deep-brain structures through surgical implantation of electrodes. It has been used as a neurosurgical intervention for various movement disorders and more recently for the treatment of severe TRD.¹ The majority of

available data on the effectiveness for DBS have been from open-label studies with reasonably long-term follow-up from 1 to 6 years.⁴ More recently, a randomized controlled study of DBS has been conducted in patients with TRD.⁵ Patients were randomized to active versus sham DBS treatment for 16 weeks (blinded) followed by an open-label continuation phase. No significant difference was observed in response rates defined as a 50% reduction in depression rating scale scores. Although the study was designed to include 208 subjects based on statistical power considerations, the study sponsor unblinded and analyzed the data after only 30 subjects completed the blinded phase. Given the small number of subjects available for these studies, improvements in both the quality and frequency of measurement of the severity of depression can lead to increases in statistical power and precision of statistical estimates.

Computerized adaptive measurement

Classical and IRT methods of measurement differ markedly in the ways in which items are administered and scored. In classical test theory, a specific counting operation measures ability (impairment), the simple sum of the individual item responses (for example, number correct). All items (symptoms) are treated as if they are equally difficult (severe). In IRT, items (symptoms) are arranged on a continuum at certain fixed points of increasing difficulty (severity). This ordering is produced by estimating the

¹Department of Neurosurgery, Rush University, Chicago, IL, USA; ²Department of Psychiatry, Rush University, Chicago, IL, USA; ³Adaptive Testing Technologies, University of Chicago, Chicago, IL, USA and ⁴Department of Medicine, Department of Public Health Sciences, University of Chicago, Chicago, IL, USA. Correspondence: Professor RD Gibbons, Department of Medicine, Department of Public Health Sciences, University of Chicago, 5841 S Maryland Avenue, W260, MC2000, Chicago, IL 60637, USA. E-mail: rdg@uchicago.edu

Received 9 May 2017; accepted 15 May 2017

parameters of an underlying model of measurement that describe how well each item discriminates between low and high levels of the underlying trait or construct (for example, depression) and how difficult (severe) the item is on the underlying continuum. Ability (impairment) is measured by the location on the continuum corresponding to the difficulty (severity) of the most difficult item answered correctly or most severe symptom expressed. In IRT, ability (severity) is measured by a scale point, not by a numerical count.

These two theories of measurement contrast sharply: if items (symptoms) are arbitrarily added or removed, the scores on traditional tests cannot be compared: scores lose their comparability if item composition is changed. The same is not true, however, of IRT scoring. If the difficulty (severity) of the items (symptoms) is changed, or a certain number of items is arbitrarily added, deleted or replaced, we do not lose comparability of scores on the scale. Only the precision of measurement at some points on the scale is affected. This property of scaled measurement, as opposed to counts of events, is the most salient advantage of IRT over classical methods of educational and psychological measurement.

CAT takes advantage of the scaled property of measurement inherent in IRT by adaptively administering a subset of items (symptoms) drawn from a much larger 'bank' of items, tailored to the specific level of ability (impairment) of each subject. After each item is administered, a provisional score and its uncertainty are computed and, on the basis of the score, the next most informative item in the bank that remains is administered. The process continues until the uncertainty falls below a predefined threshold. The paradigm shift is from short fixed-length tests with varying precision across individuals to tests with fixed-precision and varying number of items. In fact, we can markedly increase precision yet minimize patient burden and completely eliminate clinician burden using CAT.

Most applications of IRT are based on unidimensional models that assume that all of the associations between the items are explained by a single primary latent dimension or factor (for example, mathematical ability). However, mental health constructs are inherently multidimensional, where, for example, in the area of depression items may be sampled from mood, cognition, behavior and somatic subdomains, which produce residual associations between items within the subdomains that are not accounted for by the primary dimension. If we attempt to fit such data to a traditional unidimensional IRT model, we will typically have to discard the majority of candidate items to achieve a reasonable fit of the model to the data. By contrast, the bi-factor IRT model⁶ permits each item to tap the primary dimension of interest (for example, depression) and one subdomain (for example, somatic complaints), thereby accommodating the residual dependence and allowing for the retention of the majority of the items in the final model. A bi-factor IRT-informed CAT approach is highly relevant for assessing depression because of its inherent multidimensionality.

The depression module of the CAT-MH (CAT-Depression Inventory, CAT-DI) is the first CAT based on multidimensional IRT.² It extracts the information from a bank of 389 symptoms using an average of 12 adaptively selected items in ~2 min, while maintaining a correlation of $r=0.95$ with the total 389 item bank score. The CAT-DI has been shown to have strong correlation with clinician ratings (Hamilton Depression Scales, HAM-D, $r=0.75$) and other self-report scales (Patient Health Questionnaire (PHQ-9) $r=0.81$; Center for Epidemiologic Studies Depression Scale (CES-D) $r=0.84$) and the results of Structured Clinical Interview for Diagnostic Statistical Manual (SCID DSM) diagnoses of major depressive disorder (MDD; sensitivity of 0.92 and specificity of 0.88).² The CAT-DI eliminates response bias produced by repeated administration of the same test items because the items change between repeat administrations even if the level of severity does

not. Nevertheless, the CAT-DI has been shown to have higher test-retest reliability ($r=0.92$) than conventional fixed-length tests (PHQ-9 $r=0.84$), which use the exact same set of items on test and retest.⁷

MATERIALS AND METHODS

The case

Informed consent and institutional review board approval were obtained for the publication of this case study. The patient is a 55-year-old female with depressive symptoms that date back 'as far as she can remember' who presented for an evaluation. Her first appointment in the outpatient psychiatric clinic occurred 2 months after an inpatient stay for a depressive episode. She had then undergone a course of eight bilateral electroconvulsive therapy (ECT) applications with only marginal response. At this first encounter her mood was described as 'depressed', rating it a 3 out of 10 in a scale where zero would be most depressed. She also reported low energy and sleeping 10 or more hours daily. She endorsed an intense lack of pleasure in daily activities and significant feelings of guilt about not being able to 'fight' her condition. She admitted to having intermittent suicide ideas with thoughts of overdosing on prescription medications. Anxiety and compulsive rituals were also reported on this visit. She had obsessive thoughts about cleanliness and regimented routines for showering and personal hygiene. In the past, she had periods of skin picking, but those were controlled at the time of the interview and no major scarring was noticeable. Her anxiety was described as concerns over multiple different things. It was constant in character with occasional peaks and about two panic attacks a month. She denied anticipatory anxiety over the attacks. She denied any recent instances of elevated mood, increased energy or disinhibition. The patient reported a smoking habit of one pack per day, but denied use of any other substances and abuse of prescription medications.

Her previous psychiatric history was marked by lifelong problems with depressed mood that led to six psychiatric hospitalizations for depression. She reported first being brought for psychiatric treatment for depression when she was 14 years old. There were also two suicide attempts, the first in her thirties and the second 2 years before visiting our clinic. Her most serious attempt led to a prolonged intensive care unit stay. The patient had been given a diagnosis of bipolar disorder in the past. Our interviews with her and collateral information were not able to identify discrete manic episodes. There was evidence that the patient underwent hypomanic episodes at times in her life, but none of these led to hospital admissions. Some of them were related to periods of substance abuse. There were no episodes of hallucinations or delusions. She reported a difficult childhood, but denied major traumatic events, nightmares or flashbacks related to untoward events during her upbringing.

Substance use history showed a 1-year period when she abused prescription opioids. She had been abstinent from these for a matter of years. There were also periods of intense use of alcohol in her early adulthood, but her drinking habit was less than four drinks a month at the time of her presentation and had been restricted to this quantity for decades. She maintains a habit of smoking one pack of cigarettes daily, and this has lasted for the last three decades.

Medical history was positive for obesity, seasonal allergies and gastroesophageal reflux disease.

Her psychiatric family history was marked by depression in both parents, her father's with alcohol abuse and a suicide attempt.

In social history her upbringing was described as poor. She obtained a Bachelor's Degree, had three marriages and divorced as many times. She was able to maintain jobs until her late thirties when she went on disability. The patient lives independently with close supervision from family members.

Extensive review of previous treatments showed that the patient had been on multiple psychotropic medications, invariably with poor results. Among those that were tried we were able to confirm the following: Alprazolam, Amitriptyline, Bupropion, Buspirone, Carbamazepine, Chlordiazepoxide, Clonazepam, Desipramine, Dextroamphetamine, Dextroamphetamine/Amphetamine, Diazepam, Divalproate, Doxepin, Fluoxetine, Gabapentin, Haloperidol, Imipramine, Isocarboxazid, Lamotrigine, Lisdexamfetamine, Lithium, Lorazepam, Methylphenidate, Mirtazapine, Oxcarbazepine, Paroxetine, Phenelzine, Quetiapine, Risperidone, Sertraline, Tranylcypromine, Trazodone, Venlafaxine, Vortioxetine and Ziprasidone. The patient also underwent different modalities of psychotherapy during her lifetime including cognitive behavior therapy. At the time of her

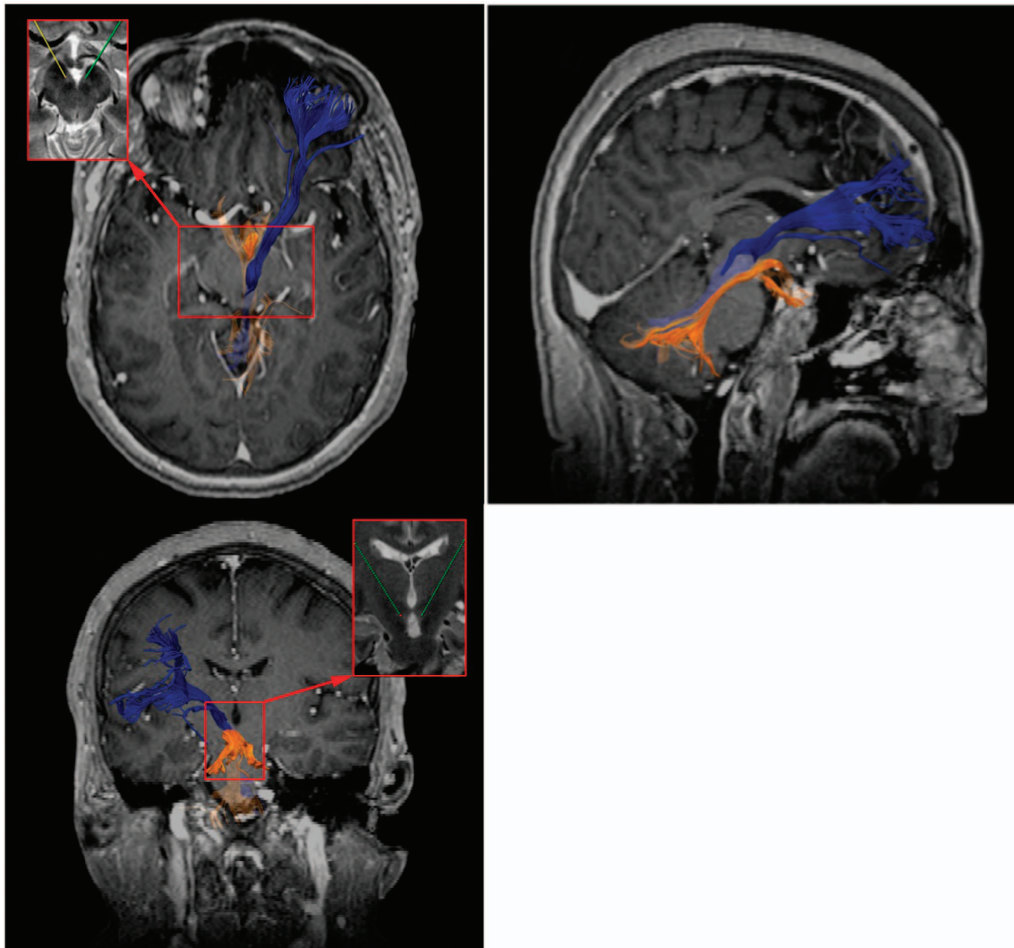


Figure 1. T1 weighted magnetic resonance imaging demonstrating the tractography of the sIMFB. (a) Axial image with the left sIMFB. Inset demonstrating the location of the right and left DBS leads. (b) Sagittal image demonstrating the prefrontal projections of the sIMFB. (c) Coronal image with inset demonstrating the location and trajectory of bilateral DBS leads. DSB, deep-brain stimulation; sIMFB, superolateral branch of the medial forebrain bundle.

presentation to our clinic she was not receiving psychotherapy treatments and declined another trial of these.

The patient entered her last hospitalization on a combination of Oxcarbazepine, Fluoxetine, Dextroamphetamine/Amphetamine and Clonazepam; their respective daily doses were 900, 40, 80 and 3 mg, and documentation showed that she had been on this regimen for over 6 months. Transcranial magnetic stimulation was considered, but patient and medical staff felt that bilateral ECT was preferred because of the severity of symptoms. The patient was discharged on a combination of Lurasidone, Oxcarbazepine and Dextroamphetamine/Amphetamine at the following daily doses 80, 900 and 80 mg. After 3 months on this combo there was still no significant response in depressive symptoms.

On a subsequent visit 1 week later, we further confirmed the patient's diagnosis using the Structured Clinical Interview for the DSM-IV Axis I Disorders. The interview was conducted by a board-certified psychiatrist. On this instrument the patient met diagnostic criteria for MDD, severe, without psychotic features; obsessive compulsive disorder, severe; social phobia, severe. She was not found to have personality disorders. Her total score on the Yale-Brown Obsessive Compulsive Scale was 24, with 18 on 'obsessions' and 6 on 'compulsions'. The score on the 21-item Hamilton Depression Scale (HAM-D) was 24. She also completed the CAT-DI² scoring 76.4 (severe) with a precision of 5.0, percentile among patients with a documented SCID-based DSM-5 MDD diagnosis of 83.2% (severe) and a probability of MDD of 0.995. This evaluation took place 2 months before DBS implantation.

We deemed the patient's depression treatment-resistant because the three following criteria were met: failed adequate trials of three different antidepressant classes; failed at least two trials of augmentation with other agents; and failed a trial of a course of ECT with at least six bilateral

sessions. On her last assessment 2 weeks before surgery, she scored 28 on the HAM-D, and CAT-DI scores were 82.3 (severe) with 4.9 precision and 99.8% probability of MDD.

Surgical planning/procedure

After a multidisciplinary review, given the refractory nature of the patient's depression, she was deemed an acceptable candidate for DBS surgery. The superolateral branch of the medial forebrain bundle (sIMFB) was selected as the target because of reported rapid symptom improvement for TRD in two recent case series. Six out of seven patients with stimulation of the sIMFB in a recent case series⁸ showed a marked resolution of symptoms. Similarly, interim results of an ongoing study have revealed three out of four patients responding markedly to DBS of the sIMFB.⁹ Stimulation of the sIMFB, a white matter pathway interconnecting various centers of the limbic system including the nucleus accumbens (NAcc), ventral tegmental area, hypothalamus and amygdala, has also been shown to occur at lower parameters than structures such as the NAcc or ventral capsule/ventral striatum.^{10–13} Because the sIMFB is not readily visualized on standard magnetic resonance imaging sequences, diffusion tensor imaging-based tractography was performed in order to visualize the sIMFB in a manner described by Anthoner *et al.*¹³ (see Figure 1).

A frame-based stereotactic technique with microelectrode recordings for target confirmation in the awake state was employed. The sIMFB was visualized bilaterally using diffusion tensor imaging tractography with direct targeting using the targeting software (FrameLink 5.1, Medtronic, Minneapolis, MN, USA). The mid-commissural coordinate system was used. Planned and final target coordinates are shown in Table 1. Microelectrode recording and microstimulation confirmed target accuracy and bilateral

DBS leads were implanted. Targeting accuracy was further confirmed by performing intraoperative computerized tomography (CT) and calculating the target coordinates.

DBS programming

The DBS leads were programmed 4 weeks after implantation to allow for resolution of brain edema. The neurosurgeon (SS) was responsible for device programming and was the only person aware of the on-off status or parameters during the blinded phase.

After the 4-week period of rest, device programming ensued and stimulation parameters were revised based on patient feedback for a period of 2 weeks, during which time optimal settings were determined. Our stimulation parameters were based on the programming foundation of movement disorder DBS patients and previous publications.¹⁴ After selection of the pulse width and frequency, each contact of each hemisphere was individually tested at various current settings, and notations on mood, anxiety and side effects were made. If suitable, incorporation of two or more contacts was made to further optimize clinical results. Of note, consistent overwhelming feelings of anxiety along with flushing and perspirations were noted during stimulation testing of all contacts at amplitudes greater than 4. We also noted gait changes with increased rigidity when stimulation settings were increased to values above 5 v. This was presumed to be related to inadvertent activation of the adjacent fibers in the fields of Forel, leading to a Parkinson-like phenotype that was reversible when the stimulation was turned off.

Subsequent to the 2-week optimization period and determination of the optimal contact(s) and settings, the device was turned off for 6 weeks to ‘wash out’ any stimulation benefit. Subsequently, the device was turned on ‘blindly’ at the previously determined optimal setting but at below threshold levels to avoid any clinical perception of stimulation or side effects. This was an attempt at reducing the placebo effect. The blinded phase of stimulation continued for 10 weeks, during which infrequent changes were made to the stimulation settings to maximize clinical improvement. At the conclusion of the blinded phase, further programming adjustments were made with the patient’s knowledge. Summary of stimulation settings are provided in Table 2.

Table 1. Targeting coordinates for the DBS lead in the left and right hemispheres

	AC-PC coordinates			Trajectory angles	
	X	Y	Z	Sagittal	Axial
<i>Left</i>					
Plan	-5.0	-2.0	-6.2	32	41
Final	-5.2	-0.2	-6.1	32	41
<i>Right</i>					
Plan	4.5	-2.5	-6.2	29	47
Final	5.8	-0.7	-6.5	29	47

Abbreviations: AC, anterior commissure; PC, posterior commissure.

Table 2. DBS program setting for the left and right hemispheres

Date	DBS settings					
	L			R		
	Current (v)	Pulse width (ms)	Frequency (Hz)	Current (v)	Pulse width (ms)	Frequency (Hz)
Week 12	2.0	60	140	2.0	60	140
Week 18	3.5	60	140	2.5	60	140
Week 20	1.5	90	140	1.5	90	140
Week 22	1.5	150	140	1.5	150	140

Abbreviations: DSB, deep-brain stimulation; Hz, Hertz; ms, millisecond; V, voltage.

Depression severity assessment

The depression module of the CAT-MH (CAT-Depression Inventory—CAT-DI) was used for daily assessments before, during and after the DBS procedure. Before and during the surgery the CAT-DI was administered by the psychiatrist attending (JB) to the patient (following electrode insertion, both before and after the stimulation was applied through the electrode). Following the surgery, a daily e-mail was sent to the patient, providing a direct internet link to the CAT-DI test. CAT-DI testing continued for 6 months postoperatively. The psychiatry team (blinded to stimulation status) met with the patient monthly following the DBS surgery and the 21-item HAM-D was administered on each visit.

RESULTS

Before turning on the electrode the patient measured 66.2, which signified moderately severe symptoms of depression. After turning on the electrode, she seemed giddy and measured 28.9, which signified normal. The patient responded to the request for CAT-DI testing 93 times out of the 193-day postoperative period (48%). The average number of measurements per week was 3.37, with an average of 2.12 days between completed interviews (maximum = 6 days). Overall, 81% of measurements were in the evening, 18% in the afternoon and 1% in the morning. The testing sessions required an average of 1 min and 55 s (median 1 min and 11 s) to complete with an average of 11.51 items (median of 11 items) being administered adaptively from the item bank.

Figure 2 presents the temporal pattern of the patient’s depressive severity scores (on a 100 point scale measured with five points of precision) before the surgery, during the surgery and daily through 6 months of follow-up. During the procedure and period following the surgery before electrical stimulation was initiated, the patient’s level of severity was markedly reduced. However, when the electrical stimulation was initiated her depressive severity was at its lowest level, but gradually increased until the stimulation was turned off and her depression decreased in severity. To this point the patient was not blinded to the status of her DBS. On week 13 her DBS was turned on for the first time at the previously determined optimal setting without the patient’s knowledge with little effect. The DBS settings were changed on week 19 and then again on week 21 without the patient’s knowledge. On week 22 the settings were changed with the patient no longer blinded. The trend in improvement continued, despite 1 day of increased severity. These changes were associated with reduced levels of depressive severity from severe (over 75) to mild (around 60), which is a difference three times the size of the uncertainty level.

The patient’s depression severity was also rated by a trained clinician using the HAM-D at baseline and at two occasions during the treatment period. The HAM-D scores are in a different metric, but show an overall decrease in severity of depression, which is similar to CAT-DI scores during those three general time periods. However, the marked fluctuations in depressive severity identified by the CAT-DI during daily monitoring are all but completely

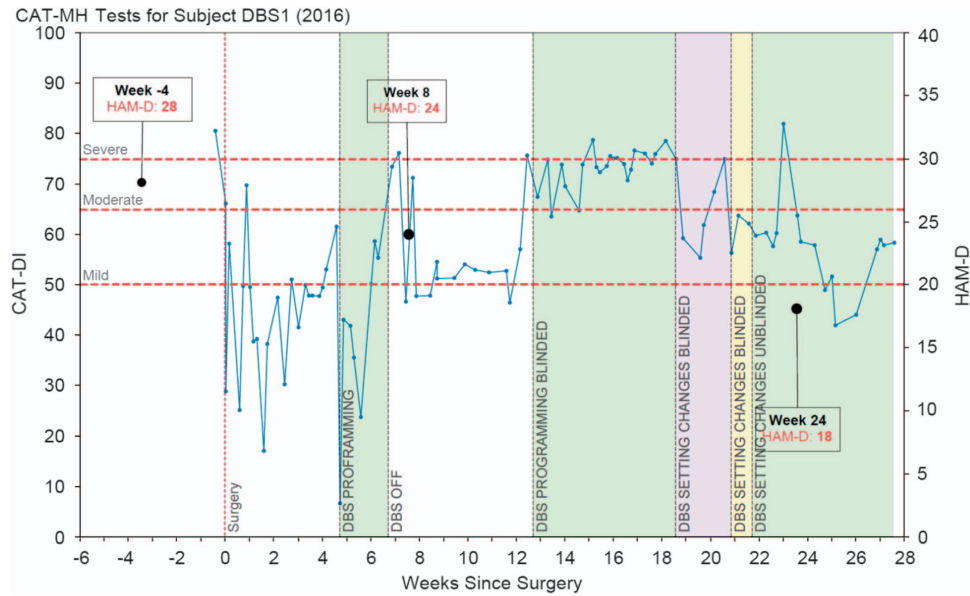


Figure 2. High-frequency depression severity measurement.

masked by the simple linear time trend formed by the three HAM-D scores.

DISCUSSION

This study supports the efficacy of CAT-DI in providing high-resolution outcome data in a TRD patient treated with DBS. Of equal importance is the demonstrated feasibility of daily depressive severity measurement at high levels of precision, equivalent to what would have required hours of assessment time each day. Compliance was high with 48% of daily e-mail prompts resulting in the patient completing a CAT-DI test (93 completed assessments over 6 months; 33.37 average number of weekly measurements; and average gap between measurements of 2.12 days). The majority (99%) of measurements were in the afternoon and evening, thereby minimizing diurnal variation. Clinician measurements of HAM-D scores confirm the general pattern of treatment response, but mask the marked variability in mood and more marked periods of benefit and decline.

Limitations of this report include the presentation of a single case with a relatively short follow-up time. Although the feasibility of CAT-DI was demonstrated, further studies including a larger cohort and long-term follow-up data are needed to determine whether our results can be generalized. Lastly, although the clinical outcome after DBS was equivocal at 6 months' follow-up in the presented case, the significance of this report lies in the ability of CAT-DI to provide a high-frequency representation of daily changes that would otherwise not be possible with in-office clinician measurements or traditional fixed-length depression self-rating scales.

There are numerous applications of this technology for high-frequency mental health measurement. These adaptive tests are available for anxiety,¹⁵ mania/hypomania¹⁶ and suicidality.¹⁷ New drugs such as ketamine, which have been used for rapid resolution of depression and suicide risk,^{18–20} require precise measurements of depressive severity and suicidal ideation repeatedly within short time intervals (for example, every 30 min), which are not possible with traditional mental health measurement systems due to response bias produced by repeated administration of the same items. Ecological momentary assessment,²¹ which are typically restricted to one or two general mood questions delivered via smartphone at random or event related intervals several times a day for weeks or months, can be

conducted with much higher precision and accuracy using CAT at any sampling interval.

CONFLICT OF INTEREST

RDG is a founder of Adaptive Testing Technologies and YC is the company's Chief Executive Officer. Adaptive Testing Technologies distributes the CAT-MH. The terms of RDG's arrangement have been reviewed and approved by the University of Chicago in accordance with its conflict of interest policies. The remaining authors declare no conflict of interest.

ACKNOWLEDGMENTS

This study was supported by funding from the National Institutes of Mental Health (NIMH) for Computerized Adaptive Testing—Depression Inventory (Grant no.: R01MH66302).

REFERENCES

- 1 Mayberg HS, Lozano AM, Voon V, McNeely HE, Seminowicz D, Hamani C et al. Deep brain stimulation for treatment-resistant depression. *Neuron* 2005; **45**: 651–660.
- 2 Gibbons RD, Weiss DJ, Pilkonis PA, Frank E, Moore T, Kim JB et al. Development of a computerized adaptive test for depression. *JAMA Psychiatry* 2012; **69**: 1104–1112.
- 3 Gibbons RD. Computerized adaptive diagnosis and testing of mental health disorders. *Annu Rev Clin Psychol* 2016; **12**: 83–104.
- 4 Malone DA Jr, Dougherty DD, Rezaei AR, Carpenter LL, Friehs GM, Eskandar EN et al. Deep brain stimulation of the ventral capsule/ventral striatum for treatment-resistant depression. *Biol Psychiatry* 2009; **65**: 267–275.
- 5 Dougherty DD, Rezaei AR, Carpenter LL, Howland RH, Bhati MT, O'Reardon JP et al. A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression. *Biol Psychiatry* 2015; **78**: 240–248.
- 6 Gibbons RD, Hedeker DR. Full-information item bi-factor analysis. *Psychometrika* 1992; **57**: 423–436.
- 7 Beiser D, Vu M, Gibbons RD. Test-retest reliability of a computerized adaptive depression test. *Psychiatr Serv* 2016; **67**: 1039–1041.
- 8 Schlaepfer TE, Bewernick BH, Kayser S, Mädlar B, Coenen VA. Rapid effects of deep brain stimulation for treatment-resistant major depression. *Biol Psychiatry* 2013; **73**: 1204–1212.
- 9 Fenoy AJ, Schulz P, Selvaraj S, Burrows C, Spiker D, Cao B et al. Deep brain stimulation of the medial forebrain bundle: distinctive responses in resistant depression. *J Affect Disord* 2016; **203**: 143–151.
- 10 Coenen VA, Honey CR, Hurwitz T, Rahman AA, McMaster J, Bürgel U et al. Medial forebrain bundle stimulation as a pathophysiological mechanism for hypomania in subthalamic nucleus deep brain stimulation for Parkinson's disease. *Neurosurgery* 2009; **64**: 1106–1114, discussion 1114–1115.

- 11 Schoene-Bake JC, Parpaley Y, Weber B, Panksepp J, Hurwitz TA, Coenen VA. Tractographic analysis of historical lesion surgery for depression. *Neuropsychopharmacology* 2010; **35**: 2553–2563.
- 12 Coenen VA, Schlaepfer TE, Maedler B, Panksepp J. Cross-species affective functions of the medial forebrain bundle-implications for the treatment of affective pain and depression in humans. *Neurosci Biobehav Rev* 2011; **35**: 1971–1981.
- 13 Anthofer JM, Steib K, Fellner C, Lange M, Brawanski A, Schlaier J. DTI-based deterministic fibre tracking of the medial forebrain bundle. *Acta Neurochir* 2015; **157**: 469–477.
- 14 Bewernick BH, Kayser S, Gippert SM, Switala C, Coenen VA, Schlaepfer TE. Deep brain stimulation to the medial forebrain bundle for depression- long-term outcomes and a novel data analysis strategy. *Brain Stimul* 2017; **10**: 664–671, pii: S1935-861X(17)30603-4.
- 15 Gibbons RD, Weiss DJ, Pilkonis PA, Frank E, Moore T, Kim JB et al. Development of the CAT-ANX: A computerized adaptive test for anxiety. *Am J Psychiatry* 2014; **171**: 187–194.
- 16 Achtyes ED, Halstead S, Smart L, Moore T, Frank E, Kupfer D et al. Validation of computerized adaptive testing in an outpatient non-academic setting. *Psychiatr Serv* 2015; **66**: 1091–1096.
- 17 Gibbons RD, Kupfer DJ, Frank E, Moore T, Beiser DG, Boudreaux ED. Development of a computerized adaptive suicide scale – The CAT-SS. *J Clin Psychiatry* (e-pub ahead of print).
- 18 Kirby T. Ketamine for depression: the highs and lows. *Lancet Psychiatry* 2015; **2**: 783–784.
- 19 Larkin GL, Beautrais AL. A preliminary naturalistic study of low-dose ketamine for depression and suicidal ideation in the emergency department. *Int J Neuropsychopharmacol* 2011; **14**: 1127–1131.
- 20 Ballard ED, Ionescu DF, Vande Voort JL, Niciu MJ, Richards EM, Luckenbaugh DA et al. Improvement in suicidal ideation after ketamine infusion: relationship to reductions in depression and anxiety. *J Psychiatr Res* 2014; **58**: 161–166.
- 21 Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008; **4**: 1–32.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

© The Author(s) 2017