

Development of the National Network of Depression Centers Mood Outcomes

Program: A Multi-site Platform for Measurement-based Care

Peter P. Zandi, Ph.D.¹, Yu-Hsun Wang, M.D.¹, Paresh Patel, M.D.¹, David Katzelnick, M.D.¹, Carolyn Turvey, Ph.D.¹, Jesse H Wright, M.D.,Ph.D.¹, Olu Aijilore, M.D., Ph.D.¹, William Coryell, M.D.¹, Christopher D. Schneck, M.D.¹, Constance Guille, M.D.¹, Erika Saunders, M.D.¹, Sophie Lazarus, Ph.D.¹, Valeria Cuellar, M.D.¹, Sudhaker Selvaraj, M.D.¹, Patricia Dill Rinvelt, M.B.A.¹, John Greden, M.D.¹, J. Raymond DePaulo, M.D.¹

¹Johns Hopkins University, Baltimore, MD (Zandi, Wang, DePaulo); University of Michigan, Ann Arbor, MI (Patel, Greden); Mayo Clinic, Rochester, MN (Katzelnick); University of Iowa Healthcare, Iowa City, IA (Turvey, Coryell); University of Louisville, Louisville, KY (Wright); University of Illinois at Chicago, Chicago, IL (Aijilore); University of Colorado, Boulder, CO (Schneck); Medical University of South Carolina, Charleston, SC (Guille); Penn State Hershey Medical Center, Hershey, PA (Saunders); Ohio State University, Columbus, OH (Lazarus); University of Texas Health Science Center, Houston, TX (Cuellar Selvaraj); National Network of Mood Disorders, Ann Arbor, MI (Dill Rinvelt);

ABSTRACT

Objectives: Mood disorders are among the most burdensome public health concerns. The National Network of Depression Centers (NNDC) is a non-profit consortium of 26 leading clinical and academic member centers around the U.S. providing care for patients with mood disorders, including depression and bipolar disorder. The NNDC has established a measurement-based care program called the Mood Outcomes Program whereby participating sites follow a standard protocol to electronically collect patient reported outcome assessments on depression, anxiety and suicidal ideation in routine clinical care. The purpose of this paper is to describe the approaches taken to develop and implement the program.

Methods: Since 2015, eight pilot sites have implemented the program and followed over 10,000 patients. This study presents descriptive statistics based on the first 24-month period of data collection.

Results: In this sample, 55% of bipolar patients (N=849) and 65% of unipolar patients (N=3,998) remained symptomatic at follow-up. Suicidal ideation, both lifetime and current, was high ranging from 27% for unipolar patients to 33% for bipolar patients. Men, unmarried individuals, and those with comorbid anxiety had a poorer longitudinal course. This initial snapshot of clinical burden is consistent with public health data indicating that mood disorders are severely debilitating.

Conclusion: This study demonstrates the potential of the Mood Outcomes Program that will be further realized as the program expands in reach and scope across additional NNDC sites with the goal of creating a nation-wide “learning health system” for mood disorders.

INTRODUCTION

Mood disorders are among the most burdensome public health concerns in the world. The lifetime prevalence of major depression in the United States is 13-17%^{1,2}, and the World Health Organization (WHO) has ranked major depression among the *most* disabling disorders on the globe³. Depression is strongly associated with suicide, particularly in populations at high risk such as veterans, the elderly and those struggling with chronic pain^{4,5}. Management of depression in medical specialty care is also critical as it is strongly associated with morbidity and non-suicide mortality in chronic conditions such as heart disease and diabetes^{6,7}. Mrazek et al.⁸ estimated that treatment resistant depression occurs in 12 to 20% of cases and incurs an annual added societal cost of \$29 to \$48 billion dollars.

There has been progress during the past five decades in improving pharmacologic and psychotherapy treatments of mood disorders. However, despite these gains, the etiology of these disorders remains poorly understood. As a result, patients and providers must often work through multiple medication trials before finding an effective treatment and achieving remission. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial found that only 37% of patients responded to the first antidepressant prescribed, and up to four medication changes were needed to achieve 67% response rates overall⁹⁻¹¹. Sustaining wellness then becomes an incremental challenge.

New approaches are needed to develop personalized treatments that are tailored to the specific needs of individual patients and can achieve more effective and timely remission and maintenance of wellness. Such precision treatments require improved understanding of the multifactorial etiologies and longitudinal course of these complex disorders. To help meet this challenge, large, well-characterized samples of patients with extended follow-up over time are needed to identify more meaningful clinical phenotypes and their underlying biological substrates. Within the United States, the medical communities surrounding cancer and heart disease have successfully developed longitudinal patient cohorts and registries. For example, the Framingham Heart Study¹²⁻¹⁵, the National Heart, Blood, and Lung Institute Registry¹⁶, and the National Cancer Registry Programs¹⁷ have contributed clinical insights that are transforming these formerly acute illnesses with high mortality to chronic illnesses with remarkable recovery, survivorship, and improved functioning.

Comparable efforts are needed to produce similar gains for the millions receiving treatment for mood disorders in psychiatric and primary care settings¹⁸. Nordic and Asian countries have developed longitudinal registries of psychiatric disorders, based on information collected in the context of their national health services¹⁹⁻²², but the United States does not have a comparable healthcare infrastructure, making it challenging to support similar nationwide registries. The closest the U.S. has to such registries are efforts to study clinical data from large health systems, such as the Mental Health Research Network²³ or the Veteran's Health Administration (VHA)^{5,24}. Longitudinal cohort studies of individuals recruited for research aims, such as the

Collaborative Study of Depression²⁵⁻²⁷ or the aforementioned Sequenced Treatment Alternatives to Relieve Depression (STAR*D)¹⁰ have also contributed enormous insights into mood disorders. However, these studies focus on select populations or the sample size of these studies are limited due to the logistical challenges and costs of on-going monitoring. The NIH recently initiated the “All of Us” research program²⁸ which seeks to overcome the limitation of small sample sizes by establishing a prospective cohort of over 1 million, but the depth of clinical information on mental illnesses to be gathered in this project is unclear.

In the United States, the recent proliferation of electronic health records and physician quality reporting systems tied to the Affordable Care Act (ACA) and the subsequent Medicare Access and CHIP Reauthorization Act (MACRA)²⁹ have promoted development of a nationwide information technology infrastructure. Major electronic health record (EHR) vendors support electronic capture of patient characteristics and healthcare utilization. Perhaps more important, current mature electronic health records enable electronic assessment of key patient reported outcomes to inform and help guide clinical decision-making and measurement-based care. This infrastructure has the potential to support identification and longitudinal monitoring of patients suffering from mood disorders in real-world clinical contexts. Clinicians can then learn what works best for these patients while providing care and, in iterations, continuously improve the care overall in what is referred to as a “learning health system”^{30,31}.

The National Network of Depression Centers (NNDC), founded in 2008, is a nationwide non-profit consortium of 26 leading clinical and academic member centers in the United States³². Its overarching mission is to integrate innovative research, clinical translation, education, and public policy to better diagnose and treat depression, bipolar disorder, and related mood disorders. In service of these goals, the NNDC is currently implementing the Mood Outcomes Program whereby participating sites follow a standard protocol to collect patient reported outcomes in the context of routine clinical care. This systematic data collection is intended to enable measurement-based care³³⁻³⁵ that can enhance patient and provider decision-making at the point-of-care and has been shown to be effective in improving overall patient outcomes³³. The collected data is then de-identified and gathered from multiple sites into a central repository that can support population health analytics and quality improvement initiatives to define best practices and aid development of precision health care. The repository also provides a platform for carrying out multi-site research through secondary data analyses or embedded studies such as pragmatic clinical trials. In this way, the Mood Outcomes Program will provide the crucial infrastructure for a nationwide learning health system on mood disorders that can drive sustained improvement in patient outcomes.

The purpose of this paper is to describe the development of the Mood Outcomes Program as standard of care in the NNDC and report on a pilot of the program at eight NNDC sites. This overview presents descriptive statistics based on the first 24-month period of data collection within the pilot sites. These data are presented to demonstrate the feasibility of collaborative data collection for the creation of a large clinical

repository, and the potential of the repository to support analyses that take advantage of ever more powerful data that are anticipated as the program is expanded and disseminated across all NNDC sites.

METHODS

The planning and development of the Mood Outcomes Program began in 2011 and entailed: 1) agreeing on the patient-reported outcome measures; 2) specifying clinical workflows for administering these measures; 3) building an information technology platform to collect the measures, visualize the results in real time, and share de-identified data with the national repository; and 4) establishing regulatory compliance and data use policies for continuous learning from the data. A pilot of the program was then initiated in 2015 and carried out in eight participating NNDC sites.

Patient Reported Outcome Measures

Starting in 2011, a group of 20 NNDC members conducted a series of phone conference meetings to discuss and develop consensus on a battery of longitudinal rating measures to be collected as “vital signs” for a new measurement-based care program for patients with mood disorders³⁶. This became known as the Standard Assessment Package. The aim was to develop a reliable, valid, brief, cost-free, self-rated, standardized and repeatable clinical tool, not just an extensive research battery such as employed in clinical trials. The goal was a tool that efficiently provides the most essential clinical information to inform point-of-care decisions. Initial clinical targets

deemed most essential were severity of mood symptoms for both depression and mania, presence and severity of comorbid anxiety, and suicidal ideation/risk.

Balancing these considerations, the following four measures were selected: 1) the 9-item Patient Health Questionnaire (PHQ- 9)^{37,38}; 2) the 7-item Generalized Anxiety Disorder Scale (GAD-7)³⁹ 3) the 5-item Altman Self-Rating Mania Scale (ASRM)⁴⁰ and 4) the 7-item patient-rated screener version of the Columbia-Suicide Severity Rating Scale (C-SSRS)⁴¹, with a baseline version to assess both recent and lifetime ideation and attempts and a follow-up version to assess ideation and attempts since the last clinic visit.

Clinical Workflows

The goal was for patients to complete the measures in the clinic waiting rooms prior to their initial and follow-up appointments and for the results to be available in real-time for review with their clinicians to inform measurement-based care. The decision to collect the measures at the clinic was made for two reasons. The first was to establish a culture that emphasized the importance of routinely collecting these mental health “vital signs” at each visit as standard of care³⁶. The second was to make it easier to address concerns about suicide risk. There were clinical and legal concerns that if patients were allowed to complete the assessments before they came to the clinics, there would have to be mechanisms in place to address situations in which patients endorsed suicidal ideation or behavior on the PHQ-9 item 9 or the CSSR-S. Such mechanisms could add prohibitive burden to implementing the program. Beyond these requirements, sites

were free to adopt whatever workflows worked best for their clinics in order to maximize program adoption.

Health Information Technology Infrastructure

Central IT Implementation. The NNDC collaborated with the Altarum Institute, a non-profit health informatics organization, to develop a health IT platform (called ePro) to support the implementation of the Mood Outcomes Program. The platform includes an on-line patient portal for collecting the patient-reported outcome measures and a clinician dashboard for reviewing the longitudinal results during the clinical encounter. The on-line patient portal allows patients to enter the self-rated scales directly into the clinical management system using tablets or other similar devices in the clinic waiting rooms. Procedures for setting up the patient accounts and logging into the portal were designed to be as simple as possible to minimize demands on support staff. Per best practices, logging in by the patient requires a unique username and password. Once the patient-reported outcome data are entered by the patient into the portal, it is available in real-time for viewing by the clinician through an on-line clinician dashboard. Once in the clinician dashboard, providers can quickly navigate to the appropriate patient and view the longitudinal results of the patient reported outcome assessments in tabular form or displayed as graphs of trends over time (Supplemental **Figure 1**). The graphs include annotation to flag values of clinical concern, including if the patient endorses items suggesting suicidal risk that need to be addressed clinically. The patient reported outcome results, including individual items, can also be copied and pasted into the patient's medical record note if desired.

On the back-end, the clinical data are maintained in a multi-tenant database such that the clinical data from each participating site are virtually partitioned to ensure data security and patient confidentiality. This clinical database is hosted on the Altarum Secure Network (ASN) which is fully compliant with all HIPAA standards for a secure environment with an encrypted database. Data are encrypted “at rest” on the server - and “in flight” during transmission between the application server and the web browser. The data stored in this multi-tenant database support the clinical goals of the measurement-based care program.

The clinical data are automatically uploaded on a near real-time basis into a separate NNDC KnowledgeBase, which is a repository of the clinical data that has been de-identified to support the population health analytics and research goals of the “learning health system”. The KnowledgeBase is housed in a secure partition of the Altarum Secure Network. A dedicated clinical coordination team at Altarum is responsible for managing the loading of the clinical data into the KnowledgeBase. The clinical data are de-identified by stripping away all Patient Health Information (PHI), including the 18 “Safe Harbor” designated identifiers⁴². Clinic site and patient IDs are anonymized, and all dates are randomly shifted by plus or minus 3 weeks to preserve the day of the week. The clinical coordination team at Altarum retains the crosswalk between the anonymous clinic site and patient IDs and any personally identifying information.

Local IT Implementation. Many sites have existing local IT platforms that they would prefer to use instead of the Altarum platform. To accommodate such preferences and maximize uptake, a local IT implementation pathway was developed for these sites. A distinct advantage of the local IT implementation pathway is that clinicians do not have to bounce between dueling IT systems when caring for their patients in the program, which reduces the resistance to adoption. Under the local IT implementation pathway, sites build the assessments and graphical views of the results into their local IT platforms. The local builds are then reviewed by a Mood Outcomes Program Steering Committee to ensure the implemented measures are consistent with NNDC standards. Once approved, local sites upload their data into the NNDC KnowledgeBase on a quarterly basis through a data loader after the data has been suitably de-identified and transformed into the appropriate data model per NNDC specifications.

Given that more than two-thirds of the NNDC sites use Epic for their electronic health record, the NNDC established a collaboration with Epic to build the Mood Outcomes Program assessment package within Epic's Foundation System to allow for more efficient implementation of the program at these sites. The Mood Outcomes Program assessments can be deployed to patients either through Epic's Welcome® self-service check in and registration module or MyChart® (Epic's patient portal) and the results viewed in graphical form by clinicians in the patient synopsis view.

Regulatory Compliance and Data Use Policies

Each NNDC site participating in the program executes a Business Associate Agreements (BAA) with Altarum to manage the necessary PHI and support the clinical care functions of the program. In addition, each site consults with its IRB as required to contribute the clinical data to the NNDC KnowledgeBase for research purposes. Because the clinical data are collected as part of standard of care and are fully de-identified prior to inclusion in any research dataset, patient consent is not obtained. This approach to consent was decided upon after extensive consultation with an Advisory Board that included members of the Mood Outcomes Program Steering Committee as well as several bioethics experts, IRB representatives and patient advocates from NNDC sites participating in the pilot. The approach and the discussions were greatly influenced by recent thinking about learning health systems^{43,44}.

The NNDC has also established a Data Use Policy to govern access to the de-identified clinical data for research purposes. NNDC members from sites who have contributed clinical data to the Mood Outcomes Program may request access to the de-identified research dataset by submitting a Data Request Form that must be pre-approved by the NNDC Mood Outcomes Steering Committee. Qualified researchers outside the NNDC may also request access to the de-identified data by submitting a Data Request Form that must be approved by the NNDC Mood Outcomes Steering Committee, as well as the NNDC Executive Committee, which consists of the NNDC leadership, and, as appropriate, each member site that contributed data.

Pilot Study

Eight NNDC centers participated in the pilot of the Mood Outcomes Program. These centers were at University of Michigan; Johns Hopkins University; Mayo Clinic; The University of Iowa; University of Louisville; University of Colorado, Anschutz Medical Campus; University of Illinois Chicago; and the Medical University of South Carolina. The Mood Outcomes Program was implemented as standard of care at outpatient psychiatry clinics in each of these centers. Patients with a primary diagnosis of any mood disorder were included in the program, and in subsequent analyses they were classified as either having a unipolar depression or bipolar disorder. Patients were over the age of 18, while one center (Johns Hopkins University) additionally recruited patients younger than age 18 to pilot the program in a pediatric population.

RESULTS

The pilot of the Mood Outcomes Program at NNDC centers began in August 2015.

Figure 1 shows the rate of patient enrollment over time from initiation of the pilot to the end of December 2018 when the network reached its target of 10,000 unique patients.

Of the eight sites who contributed data, one used Epic to implement the program via the local pathway and integrated assessment into the initial evaluation of all patients in its outpatient clinics. This site contributed more than double the number of patients than the next leading site, demonstrating the advantages of the Epic implementation to facilitate adoption of the program. As a result, we anticipate more sites will adopt the Epic implementation as we continue to expand the program.

Table 1 presents the demographic characteristics and summary results on the Mood Outcomes Program measures at the baseline visit for patients enrolled over the first 24-month period of data collection broken down by bipolar disorder (N=849) versus unipolar depression (N=3998). The sample is predominantly female, reflecting closely the two-to-one gender ratio reported in both treated and community samples for unipolar mood disorders. Both diagnostic subgroups were middle-aged on average, with a mean age of 42.9 ± 15.7 for those with bipolar disorder and 40.2 ± 16.4 for those with unipolar depression.

The mean baseline PHQ-9 and GAD-7 scores fell in the mild-to-moderate range. They were 12.3 ± 7.1 and 10.5 ± 6.4 , respectively, for patients with unipolar depression, and 11.2 ± 7.2 and 9.5 ± 6.5 , respectively, for those with bipolar disorder. It is likely that few patients were in manic episodes at the time of assessment, reflected by the mean ASRM scores, which were 3.5 ± 3.6 for those with bipolar disorder and 2.7 ± 2 for unipolar depression. Nearly one third of the patients in both diagnosis subgroups had a lifetime history of planned or actual suicide attempt. Note the baseline visit for the registry does not necessarily fall at treatment initiation as many patients had been in treatment for a period before the baseline assessment.

Figures 2 reflects the longitudinal course of depression including patients who had a PHQ-9 greater than 10 at the baseline visit with a follow up visit at least 30 days after the baseline assessment. Between 55 and 60% of unipolar depressed and bipolar disorder patients remained symptomatic at their last follow-up on the PHQ-9. For

unipolar depressed patients, 16.1% achieved remission (<5 on the PHQ-9) while 18.7% of the patients with bipolar disorder achieved remission from depression.

Figure 3 presents explorations of the relation between patient characteristics – including gender, marital status, diagnosis and co-morbid anxiety – and longitudinal course of depression symptom severity. For example, with gender, the trajectory diverged between men and women patients at approximately 6-month follow-up with men showing clearly poorer recovery rates at 12- and 24-month follow-up. Patients who were married showed improved recovery at 12- and 24-months relative to all other marital categories. Though the trajectory for depression remission was largely comparable between unipolar depression and bipolar disorder patients, comorbid anxiety as indicated by a baseline GAD-7 score of 10 or higher, portended a significantly poorer course with marked divergence at 24-month follow-up.

DISCUSSION

The progress of the NNDC Mood Outcomes Program thus far demonstrates the feasibility and potential of collaborative systematic collection of patient-reported outcomes as part of a measurement-based care program for the treatment for patients with mood disorders. The program has met its initial goals of enrolling over 10,000 patients (current enrollment is at 10,570) and using standard assessments routinely collected through a flexible health informatics infrastructure to inform treatment at the point of care. As we disseminate the program to additional sites across the network and continue to follow our patients with multiple assessments, we anticipate rapidly growing

a progressively richer data set to explore the longitudinal course of mood disorders. The current pilot, therefore, provides an important step towards achieving our goal of a fully realized “learning health system”^{30,31} for mood disorders.

Preliminary analysis of the data collected in the pilot demonstrates that new and more powerful precision-driven treatments, ideally guided by biomarkers, are needed to understand and optimally treat mood disorders. In this sample, nearly 60% of the mood disorder patients remained symptomatic at follow-up. The analyses of the association between patient characteristics and longitudinal course revealed a poorer course for men, unmarried individuals, and those with comorbid anxiety. This initial snapshot of clinical burden, chronicity, and risk for suicide concurs with the converging public health data from the World Health Organization and others indicating that mood disorders are among the most debilitating health conditions^{45,46}.

These results reflect the common experience of treating mood disorders, that is, our best treatments and most skilled clinicians still only provide most patients with partial improvements in their mood disorders^{9,26}. Although this should give one pause, the results need to be interpreted with several important caveats in mind. The data are collected in a naturalistic setting as patients receive usual care. As a result, patients may enter the program at different points in their episode of care, and those who are doing better may be less likely to return to the clinic for additional treatment. Thus, with the data collected during this pilot, we may only be observing a partial snapshot of the full trajectory of illness, especially among those with more challenging outcomes. We

anticipate that as we continue to follow more patients over longer periods of time we will get a more complete picture of the different illness trajectories and we will be able to examine how differential follow-up affects the interpretation of the results.

Despite the initial excitement generated by the introduction of selective serotonin reuptake inhibitors (SSRIs) and, more recently, multiple-mechanisms antidepressants⁴⁷, or the promise of new findings in psychiatric genetics⁴⁸, our field has stalled in its ability to significantly reduce the medical burden of a large proportion of our patients suffering mood disorders. Larger, longitudinal, measurement-based data bases are essential if we are to advance to more precise and effective treatments. The current pilot illustrates that collection of such data is achievable, and it provides important lessons that can facilitate the wider implementation of this program and others like it (see **Box 1**).

Encouraged by the success of the pilot, we are currently working to expand the reach and scope of the Mood Outcomes Program. We are using the lessons of the pilot to rapidly disseminate the program to other NNDC sites and additional clinics within these sites, which should further accelerate the enrollment of patients into the program. Of particular note, we seek to expand the program into primary care where the majority of patients with mood disorders often receive care, as well as into other specialty clinics such as women's mood disorders clinics or brain stimulation clinics. In addition, we are working on plans to collect detailed clinical data extracted from participating sites' electronic medical records – such as healthcare utilization, diagnostic information, treatment histories, and laboratory results – to link with the patient-reported outcome

measures in the KnowledgeBase. We envision using the infrastructure of the Mood Outcomes Program and the rich, longitudinal data gathered in the KnowledgeBase as a platform to support large-scale, multi-site quality improvement and research efforts across the NNDC centers. This may involve using big-data analytics with the gathered data, as well conducting embedded studies in real-world settings to generate new data. The goal will be to rapidly translate what is learned from these efforts back into improved care for our patients, and thus achieve the vision of a nation-wide learning health system for mood disorders.

Author and Article Information

The Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine (Zandi, DePaulo); The Department of Mental Health, Johns Hopkins Bloomberg School of Public Health (Zandi, Wang); The Department of Psychiatry, University of Michigan (Patel); The Department of Psychiatry and Psychology, Mayo Clinic (Katzelnick); the Department of Psychiatry, University of Iowa Healthcare (Turvey); Depression Center, University of Louisville (Wright); Department of Psychiatry, University of Illinois at Chicago (Aijilore); Department of Psychiatry, University of Iowa College of Medicine (Coryell); Department of Psychiatry, University of Colorado (Schneck); Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina (Guille); Department of Psychiatry, Penn State Hershey Medical Center (Saunders); Department of Psychology, Ohio State University (Lazarus); Center of Excellence on Mood Disorders, University of Texas Health Science Center (Cuellar); Department of Psychiatry and Behavioral Sciences, University of Texas Health Science Center (Selvaraj); National Network of Mood Disorders (Dill Rinvelt); Comprehensive Depression Center, University of Michigan (Greden).

Please send all correspondence to Peter P. Zandi (pzandi1@jhu.edu); 550 North Broadway, Baltimore, MD, 21205

Acknowledgements

This work was made possible by a research collaboration supported by the National Network of Depression Centers ([NNDC](#)), an inter-dependent consortium of academic

depression centers. Operational support was provided by: Edie Douglas, MPH (Medical University of South Carolina), Kara Glazer (Johns Hopkins University), Marjorie Gresbrink (Mayo Clinic), Deanna Hofschulte, CCRP (Mayo Clinic), Dane Larsen (NNDC), Martha Shaw (University Iowa), Dana Steidtmann, PhD (University of Colorado Anschutz Medical Campus), and Carol Wahl (University of Louisville).

[Altarum](#), a nonprofit research and consulting organization that creates and implements solutions to advance health, provided technical expertise to build and maintain the Mood Outcomes system. Operational support from Altarum was provided by: Rachelle May-Gentile, Mike Grim, and Ryan Callahan.

Special thanks also go to Katie Hurtis and other members of the Epic staff for their ongoing assistance and support during this effort.

Financial Disclosure

Dr. Peter Zandi, Dr. Yu-Hsun Wang, Dr. Paresh Patel, Dr. Carolyn Turvey, Dr. Christopher Schneck, Dr. Sophie Lazerus, Dr. Valeria Cuellar, Dr. William Coryell, and Ms. Patricia Dill Rinveld do not have any financial interests to report.

Dr. David Katzelnick has commercial interests with Principle Healthcare Technology Systems

Dr. Jesse Wright has an equity interest in Empower Interactive and Mindstreet, developers and distributors of digital behavioral health programs. His conflict of interest is managed with an agreement with the University of Louisville. He receives book

royalties from American Psychiatric Press, Inc., Guilford Press, and Simon and Schuster, and he receives grant support from the Agency for Healthcare Research and Quality and the Oticon Foundation.

Dr. Olu Aijilore is the Co-founder and Chief Medical Officer of Keywise, Inc. He is on the Advisory Board of both Blueprint Health and Embodied Labs, and is a consultant for Quartet Health

Dr. Constance Guille is a member of Sage therapeutics advisory board and speakers bureau

Dr. Erica Saunders Receive an honorarium of \$2500 for work as an associate editor from the Journal of Clinical Psychiatry for 2017, 2018

Dr. Selvaraj has received speaking honoraria from Global Medical Education and honoraria from British Medical Journal Publishing Group, and owns shares at Flow Med Tech and site-investigator for a clinical trial by COMPASS Pathways Limited

Dr. Greden is a consultant for Clarigent, Genomind, Med-IQ, Inc., MoldX, Sage Therapeutics, and is involved in a pharmacogenomics study funded by AssureX and Myriad Neuroscience. He is involved commercially with Jassen Pharmaceutical, Naurex (Allergan) Pharmaceutical, Cerecor Pharmaceutical, and NeuralStem. He is a board member or does pro bono work with the National Network of Depression Centers, the Canadian Depression Research Intervention Network, Le Royal, American Foundation for Suicide Prevention, Depression and Bipolar Support Alliance, Informed Medical Decision Making Foundation, Psychiatric Research Society, and Great Lakes Bay Region

Dr. J. Raymond DePaulo is an unpaid consultant for the Assurex Health/ Myriad Company; he owns stock in CVS which now owns Aetna Insurance

REFERENCES

1. Hasin DS, Goodwin RD, Stinson FS, Grant BF. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Arch Gen Psychiatry*. 2005;62(10):1097-1106.
2. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res*. 2012;21(3):169-184.
3. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2017;390(10100):1211-1259.
4. *Suicide Among Veterans and Other Americans*. 2017.
5. Conner KR, Bohnert AS, McCarthy JF, et al. Mental disorder comorbidity and suicide among 2.96 million men receiving care in the Veterans Health Administration health system. *J Abnorm Psychol*. 2013;122(1):256-263.
6. Noel PH, Williams JW, Jr., Unutzer J, et al. Depression and comorbid illness in elderly primary care patients: impact on multiple domains of health status and well-being. *Ann Fam Med*. 2004;2(6):555-562.
7. Unutzer J, Powers D, Katon W, Langston C. From establishing an evidence-based practice to implementation in real-world settings: IMPACT as a case study. *Psychiatr Clin North Am*. 2005;28(4):1079-1092.
8. Mrazek DA, Hornberger JC, Altar CA, Degtiar I. A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013. *Psychiatr Serv*. 2014;65(8):977-987.
9. Fava M, Rush AJ, Trivedi MH, et al. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. *Psychiatr Clin North Am*. 2003;26(2):457-494, x.
10. Rush AJ, Fava M, Wisniewski SR, et al. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials*. 2004;25(1):119-142.
11. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163(11):1905-1917.
12. Dogan MV, Beach SRH, Simons RL, Lendasse A, Penaluna B, Philibert RA. Blood-Based Biomarkers for Predicting the Risk for Five-Year Incident Coronary Heart Disease in the Framingham Heart Study via Machine Learning. *Genes (Basel)*. 2018;9(12).
13. Dogan MV, Grumbach IM, Michaelson JJ, Philibert RA. Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study. *PLoS One*. 2018;13(1):e0190549.
14. Enserro DM, Vasan RS, Xanthakis V. Twenty-Year Trends in the American Heart Association Cardiovascular Health Score and Impact on Subclinical and Clinical Cardiovascular Disease: The Framingham Offspring Study. *J Am Heart Assoc*. 2018;7(11).

15. Vasani RS, Xanthakis V, Lyass A, et al. Epidemiology of Left Ventricular Systolic Dysfunction and Heart Failure in the Framingham Study: An Echocardiographic Study Over 3 Decades. *JACC Cardiovasc Imaging*. 2018;11(1):1-11.
16. Gupta N, Lee HS, Ryu JH, et al. The NHLBI LAM Registry: Prognostic Physiologic and Radiologic Biomarkers Emerge From a 15-Year Prospective Longitudinal Analysis. *Chest*. 2019;155(2):288-296.
17. Sirirungreung A, Buasom R, Jiraphongsa C, Sangrajrang S. Data Reliability and Coding Completeness of Cancer Registry Information Using Reabstracting Method in the National Cancer Institute: Thailand, 2012 to 2014. *J Glob Oncol*. 2018;4:1-9.
18. Sowa NA, Jeng P, Bauer AM, et al. Psychiatric Case Review and Treatment Intensification in Collaborative Care Management for Depression in Primary Care. *Psychiatr Serv*. 2018;69(5):549-554.
19. Kawanishi C, Ishii T, Yonemoto N, et al. Protocol for a prospective multicentre registry cohort study on suicide attempters given the assertive case management intervention after admission to an emergency department in Japan: post-ACTION-J Study (PACS). *BMJ Open*. 2018;8(9):e020517.
20. Munk-Jorgensen P, Mortensen PB. The Danish Psychiatric Central Register. *Dan Med Bull*. 1997;44(1):82-84.
21. Vernal DL, Kapoor S, Al-Jadiri A, et al. Outcome of Youth with Early-Phase Schizophrenia-Spectrum Disorders and Psychosis Not Otherwise Specified Treated with Second-Generation Antipsychotics: 12 Week Results from a Prospective, Naturalistic Cohort Study. *J Child Adolesc Psychopharmacol*. 2015;25(7):535-547.
22. Saito S, Horinouchi T, Nakagami Y, et al. Approaches to suicide prevention: Ideas and models presented by Japanese and international early career psychiatrists. *Psychiatry Clin Neurosci*. 2018;72(9):741.
23. Coleman KJ, Stewart C, Waitzfelder BE, et al. Racial-Ethnic Differences in Psychiatric Diagnoses and Treatment Across 11 Health Care Systems in the Mental Health Research Network. *Psychiatr Serv*. 2016;67(7):749-757.
24. Desai MM, Rosenheck RA, Craig TJ. Case-finding for depression among medical outpatients in the Veterans Health Administration. *Med Care*. 2006;44(2):175-181.
25. Coryell W, Endicott J, Maser JD, Mueller T, Lavori P, Keller M. The likelihood of recurrence in bipolar affective disorder: the importance of episode recency. *J Affect Disord*. 1995;33(3):201-206.
26. Coryell W, Winokur G, Solomon D, Shea T, Leon A, Keller M. Lithium and recurrence in a long-term follow-up of bipolar affective disorder. *Psychol Med*. 1997;27(2):281-289.
27. Fiedorowicz JG, Coryell WH, Rice JP, Warren LL, Haynes WG. Vasculopathy related to manic/hypomanic symptom burden and first-generation antipsychotics in a sub-sample from the collaborative depression study. *Psychother Psychosom*. 2012;81(4):235-243.
28. All of Us Research Program. <https://allofus.nih.gov>. Accessed May 18th, 2019.
29. Clough JD, McClellan M. Implementing MACRA: Implications for Physicians and for Physician Leadership. *JAMA*. 2016;315(22):2397-2398.

30. Friedman C, Rubin J, Brown J, et al. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *J Am Med Inform Assoc.* 2015;22(1):43-50.
31. Grossman C, Powers B, JM M, eds. *Digital infrastructure for the learning health system: The foundation for continuous health improvement in health and health care.* Washington, DC: National Academies Press; 2011. The foundation for continuous improvement in health and health care workshop series summary.
32. Greden JF. The National Network of Depression Centers: progress through partnership. *Depress Anxiety.* 2011;28(8):615-621.
33. Guo T, Xiang YT, Xiao L, et al. Measurement-Based Care Versus Standard Care for Major Depression: A Randomized Controlled Trial With Blind Raters. *Am J Psychiatry.* 2015;172(10):1004-1013.
34. Rush AJ. Isn't It About Time to Employ Measurement-Based Care in Practice? *Am J Psychiatry.* 2015;172(10):934-936.
35. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry.* 2006;163(1):28-40.
36. Trivedi MH, Jha MK, Kahalnik F, et al. VitalSign(6): A Primary Care First (PCP-First) Model for Universal Screening and Measurement-Based Care for Depression. *Pharmaceuticals (Basel).* 2019;12(2).
37. Thekkumpurath P, Walker J, Butcher I, et al. Screening for major depression in cancer outpatients: the diagnostic accuracy of the 9-item patient health questionnaire. *Cancer.* 2011;117(1):218-227.
38. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care.* 2003;41(11):1284-1292.
39. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-1097.
40. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. *Biol Psychiatry.* 1997;42(10):948-955.
41. Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry.* 2011;168(12):1266-1277.
42. Guidance on Satisfying the Safe Harbor Method. Department of Health and Human Services. <https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html#safeharborguidance>. Accessed October 21, 2019.
43. Faden RR, Beauchamp TL, Kass NE. Informed consent for comparative effectiveness trials. *N Engl J Med.* 2014;370(20):1959-1960.
44. Faden RR, Beauchamp TL, Kass NE. Informed consent, comparative effectiveness, and learning health care. *N Engl J Med.* 2014;370(8):766-768.
45. Edlund MJ, Wang J, Brown KG, et al. Which mental disorders are associated with the greatest impairment in functioning? *Soc Psychiatry Psychiatr Epidemiol.* 2018;53(11):1265-1276.
46. Global Burden of Disease Study C. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases

- and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;386(9995):743-800.
47. Connolly KR, Thase ME. Emerging drugs for major depressive disorder. *Expert Opin Emerg Drugs*. 2012;17(1):105-126.
 48. Howard DM, Adams MJ, Clarke TK, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*. 2019;22(3):343-352.

Table 1: Sample Description and Baseline¹ Symptom Severity²

	Bipolar Disorder (N = 849)	Unipolar Depression (N = 3,998)
Age (Mean±SD)	42.9 ± 15.7	40.2 ± 16.4
Female (N; %)	522; 61.5	2,707; 67.7
Baseline PHQ-9 (Mean±SD)	11.2 ± 7.2	12.3 ± 7.1
Baseline GAD-7 (Mean±SD)	9.5 ± 6.5	10.5 ± 6.4
Baseline ASRM (Mean±SD)	3.5 ± 3.6	2.7 ± 2.9
Positive lifetime planned or attempted suicide on baseline C-SSRS (N; %)	284; 33.5	1,103; 27.6
Positive response to item #9 on baseline PHQ-9 (N; %)	254; 29.9	1,339; 33.5

N = number; SD = standard deviation; PHQ-9 = Patient Health Questionnaire-9 (scores ranging from 0-27, higher scores indicating increased symptom severity for depression); GAD-7 = General Anxiety Disorder-7 (scores ranging from 0-21, higher scores indicating increase symptom severity for anxiety); ASRM = Altman Self-Rating Mania (scores ranging from 5-25, higher scores indicating greater severity of manic symptoms); C-SSRS = Columbia-Suicide Severity Rating Scale

¹ Baseline is defined as the first available measure for each patient after enrollment into the measurement-based care program

² Comparisons by diagnosis of all sample characteristics and baseline symptom measures were significantly different at p<0.01

Box 1. Lessons Learned

Patient Experience. Minimize barriers and friction to patient participation

Clinician Experience. Optimize the clinician's workflow by integrating into the EHR

Sustainability. Find the balance between a comprehensive vs. practical data set

Education. Patient and clinician buy-in is critical to engagement

Leadership. A local champion is essential to success

Planning. Start early to secure IT support and local governance

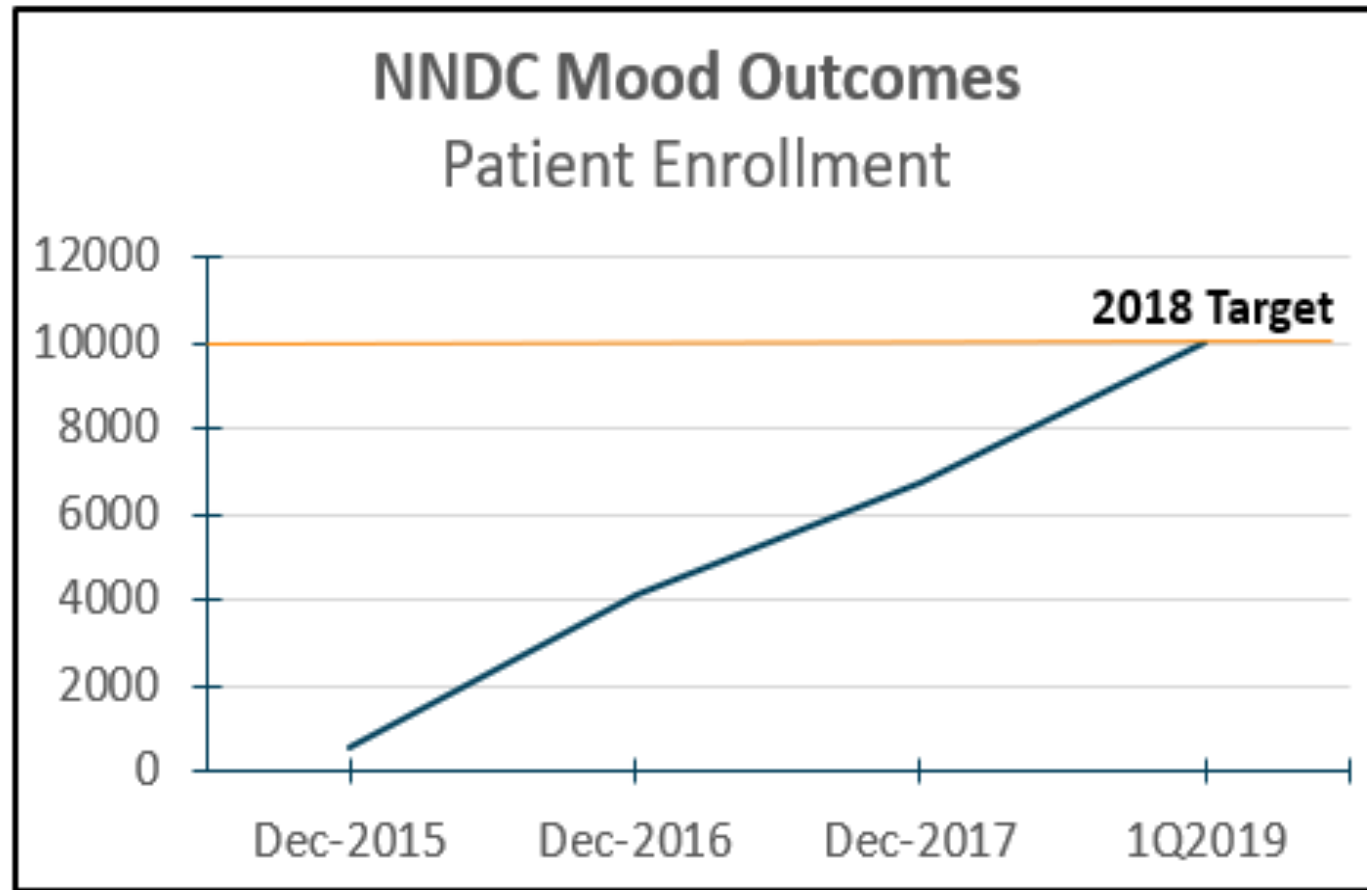


Figure 1. Patient enrollment in the Mood Outcomes Program in 8 NNDC sites between 2015 and 2018.

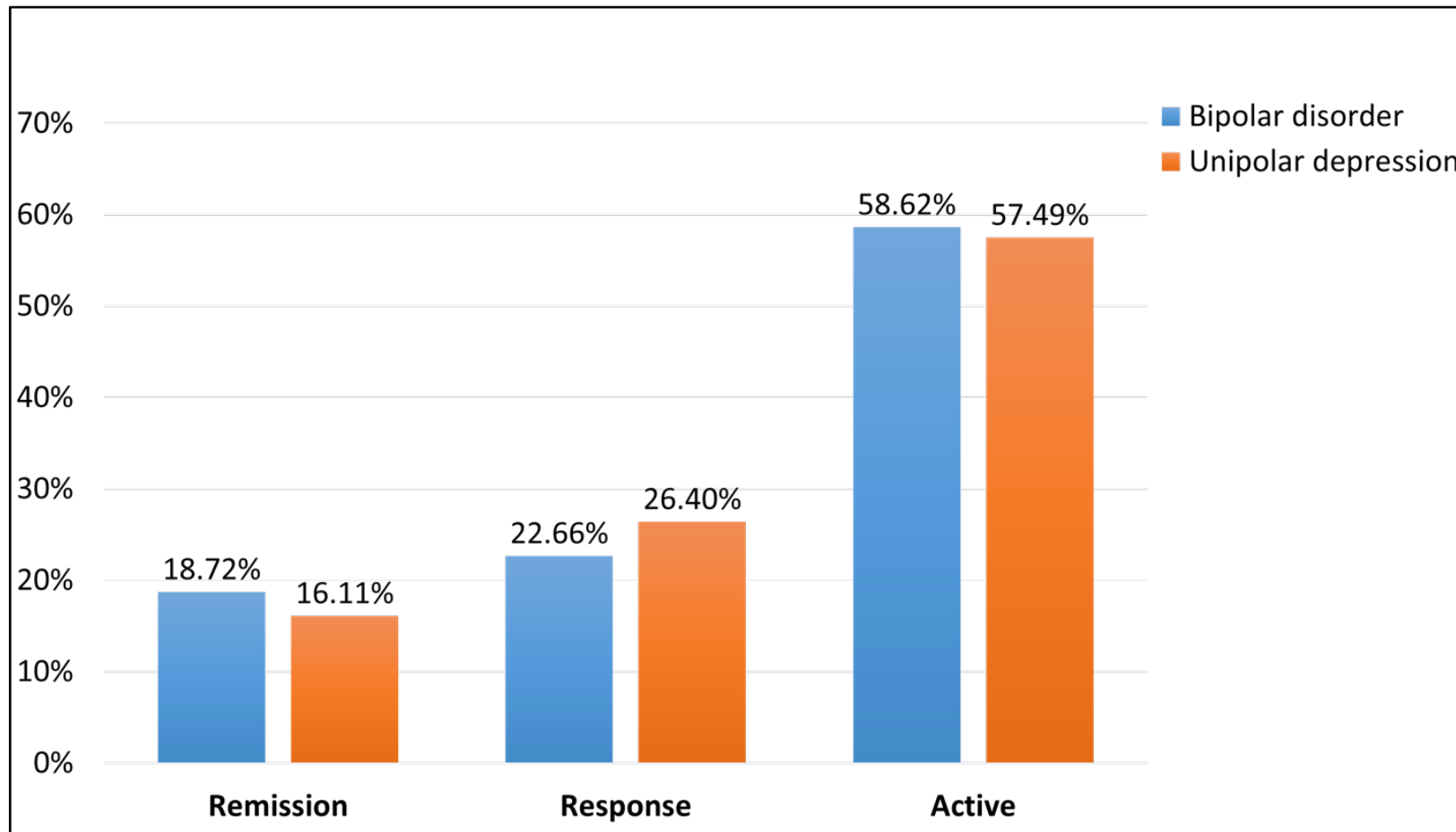


Figure 2: Status on the PHQ-9 at the last follow-up visit among patients who had either bipolar disorder or unipolar depression and entered the Mood Outcomes Program with PHQ-9 \geq 10 with at least one follow-up visit >30 days from their baseline visit. Active: PHQ-9 \geq 10; Response: PHQ-9 5-9; Remission: PHQ-9 0-4.

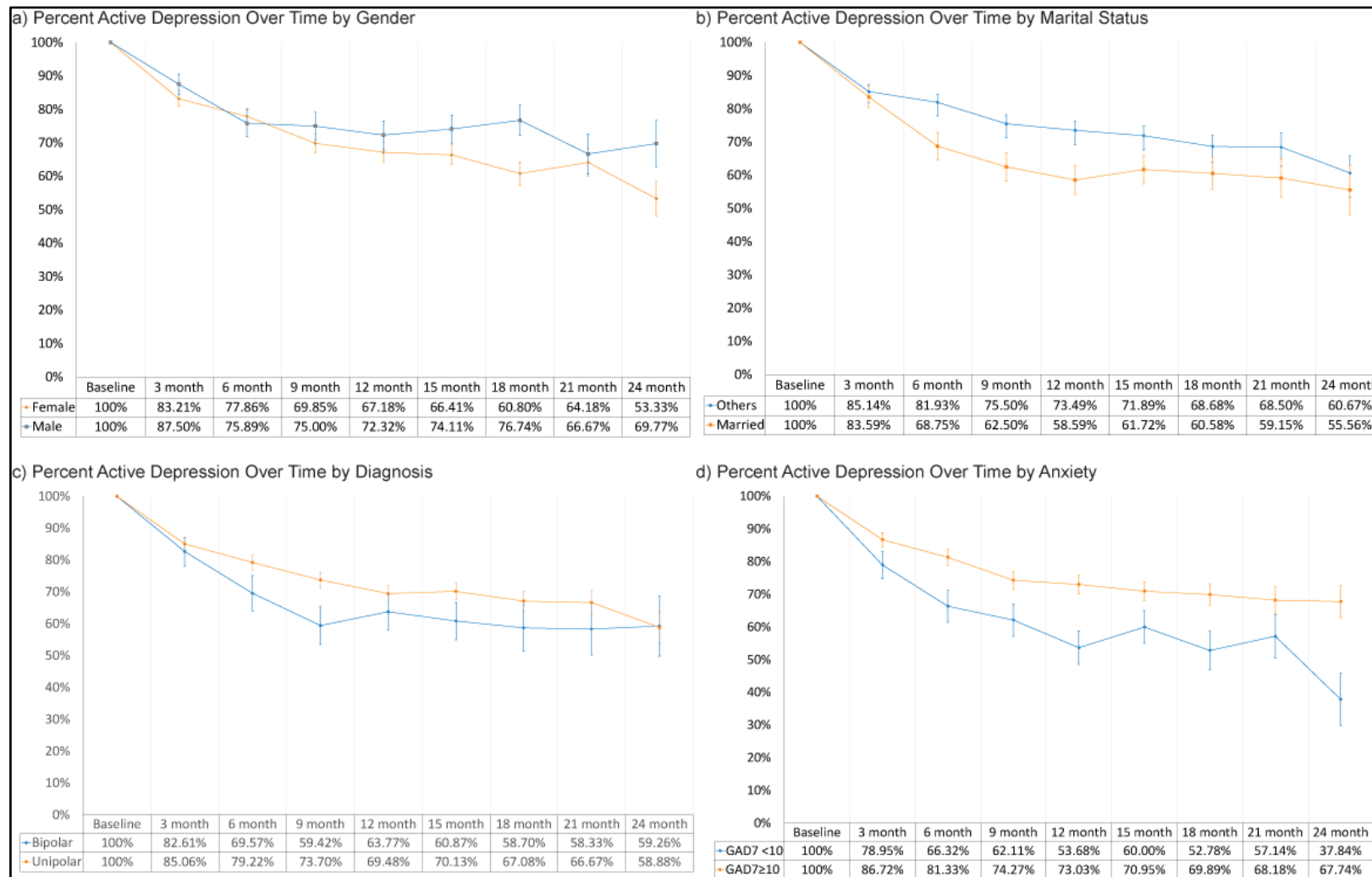
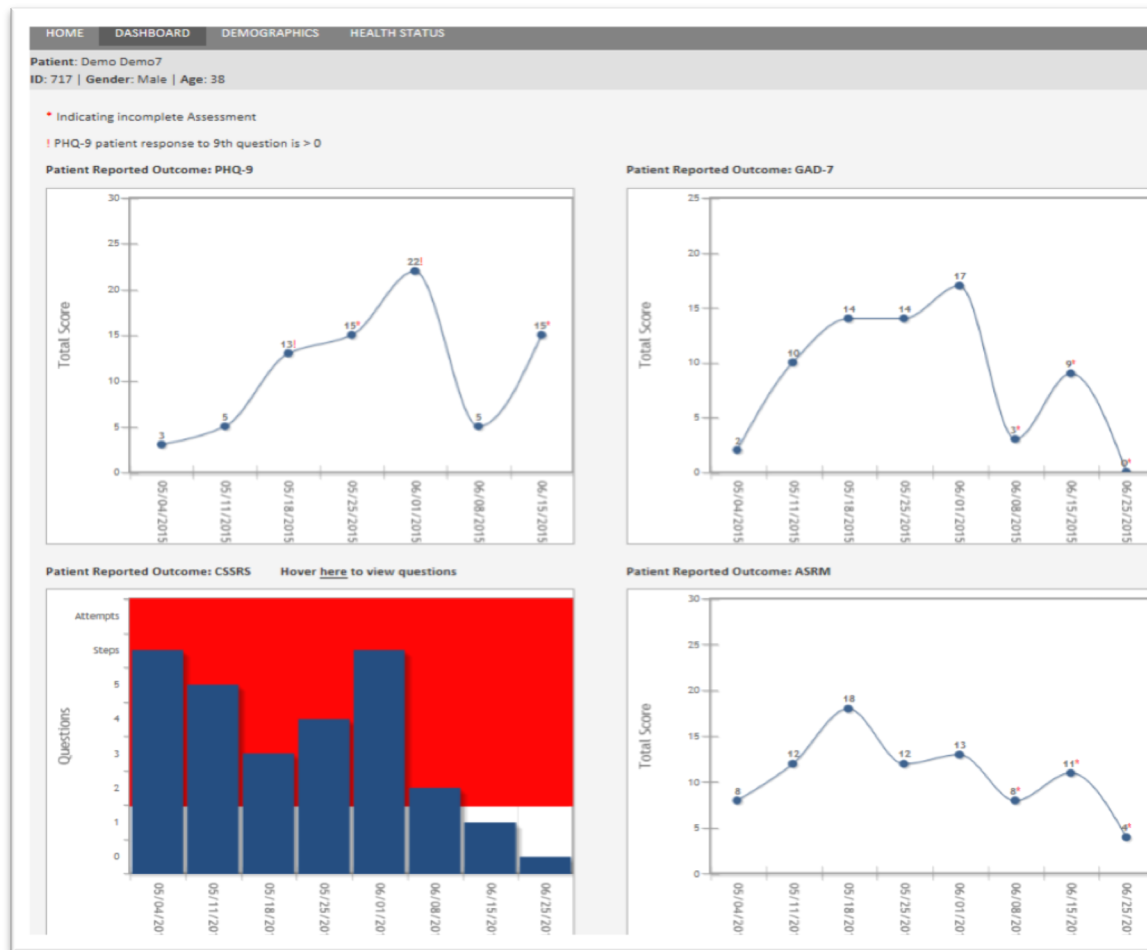


Figure 3. Percent of patients with active depression (PHQ-9>10) over time among those who entered the Mood Outcomes Program with PHQ-9_≥10 and had at least one follow-up visit >30 days from their baseline visit by: a) gender (male vs female); b) marital status (married vs other); c) diagnosis (bipolar disorder vs unipolar depression); and d) co-morbid anxiety (GAD<10 vs GAD_≥10).



Supplemental Figure 1. Mood Outcomes Program Clinician Dashboard displaying the longitudinal results of a patient on the four patient reported outcome measures: PHQ-9, GAD-7, ASRM, and C-SSRS. The responses are flagged if answers are incomplete and if there is concern about suicide risk that the providers should attend with the patient. The provider may additionally click to answers to specific questions and copy and paste the graphs and responses to questions to the patient’s electronic medical record if desired.