ORIGINAL RESEARCH

Prediction of Primary Care Depression Outcomes at Six Months: Validation of DOC-6 ©

Kurt B. Angstman, MS, MD, Gregory M. Garrison, MD, MS, Cesar A. Gonzalez, PhD, Daniel W. Cozine, BCE, Elizabeth W. Cozine, MD, and David J. Katzelnick, MD

Background: The goal of this study was to develop and validate an assessment tool for adult primary care patients diagnosed with depression to determine predictive probability of clinical outcomes at 6 months.

    Methods: We retrospectively reviewed 3096 adult patients enrolled in collaborative care management (CCM) for depression. Patients enrolled on or before December 31, 2013, served as the training set (n = 2525), whereas those enrolled after that date served as the preliminary validation set (n = 571).

    Results: Six variables (2 demographic and 4 clinical) were statistically significant in determining clinical outcomes. Using the validation data set, the remission classifier produced the receiver operating characteristics (ROC) curve with a c-statistic or area under the curve (AUC) of 0.62 with predicted probabilities than ranged from 14.5% to 79.1%, with a median of 50.6%. The persistent depressive symptoms (PDS) classifier produced an ROC curve with a c-statistic or AUC of 0.67 and predicted probabilities that ranged from 5.5% to 73.1%, with a median of 23.5%.

    Conclusions: We were able to identify readily available variables and then validated these in the prediction of depression remission and PDS at 6 months. The DOC-6 tool may be used to predict which patients may be at risk for worse outcomes. (J Am Board Fam Med 2017;30:281–287.)

Keywords: Depression, Primary Health Care, Retrospective Studies, ROC Curve

Depressive symptomatology is common among patients seeking primary care. The implementation of collaborative care models (CCMs) has contributed to addressing and improving clinical outcomes for depressive disorders in primary care settings.[1–3] CCMs provide additional resources that otherwise would not be available, such as care managers and psychiatric oversight to adapt treatment recommendations. Despite improved outcomes for depression, we previously demonstrated in our CCM population that the predicted probability of persistent depressive symptoms (PDSs) 6 months after implementing the CCM ranged from 6.6% to 42.1%.[4] The predictive factors involved were clinical diagnosis (first episode vs recurrent major depression), depression severity, anxiety severity, and the result of the Mood Disorder Questionnaire (MDQ[5]) screen. Furthermore, prior studies at our institution associated a higher baseline severity of depressive symptoms (as measured on the 9-item Patient Health Questionnaire [PHQ-9[6]]) and anxiety (as measured by the 7-item Generalized Anxiety Disorder Questionnaire [GAD-7[7]]), and a positive MDQ screen with worse clinical outcomes in a CCM.[8–11]

    In addition to mental health comorbidities serving as risk factors for poor clinical outcomes of depressive symptoms in primary care, psychosocial difficulties are predictive of poor outcomes for the treatment of depression in primary care.[12] More-

over, psychosocial difficulties are associated with complications with treatment engagement and negatively affect medical outcomes in primary care.[12–18] In Europe, the PREDICT study developed a multifactor risk score for predicting the development of depression among primary care patients who were asymptomatic for depression.[19,20] To our knowledge, however, no empirically derived tool is available to predict depression outcomes among patients in primary care who have recently been diagnosed with major depression.

Given the rate of PDSs among patients in primary care and the risk factors associated with poor clinical outcomes for the treatment of depression in primary care, practical and clinical implications exist for predicting depression outcomes. By estimating the probability of remission from depressive symptomatology using readily available information based on patients' demographic and clinical characteristics, clinicians may be able to optimize outcomes by adapting treatment recommendations, specifically for patients identified to be at higher risk for poor outcomes.

The goal of this study was to develop and validate an assessment tool for adult primary care patients diagnosed with depression that used their initial clinical and demographic characteristics by modeling the probabilities of remission and PDSs at 6 months. By estimating the relative risk associated with the clinical and demographic patient variables, clinicians may be able to adapt therapies, specifically for patients at higher risk for PDSs.

Methods

This study was approved by the Mayo Clinic Institutional Review Board. Adult patients in this study were identified from the primary care practices at the Mayo Clinic in southeastern Minnesota (empaneled patients from the Department of Family Medicine, Division of Primary Care Internal Medicine, and Division of Community Pediatrics and Adolescent Medicine). A CCM for depression care was implemented at 2 sites in 2008 and expanded to all 5 clinical sites (~100,000 adult patients) by March 2010. Our CCM, through assistance from a specifically trained registered nurse care manager, with weekly oversight of the patients by a psychiatrist, included a depression registry, treatment guidelines, and relapse prevention counseling, and has been reviewed previously.[21,22] All adult patients

diagnosed (by primary or specialty care clinicians) with major depression or dysthymia and with an initial PHQ-9 score ≥10 had the option of enrolling in the program. The only exclusionary criterion for entry into the CCM was a clinical diagnosis of bipolar disorder (an abnormal MDQ screen was acceptable). Patients who were diagnosed with major depressive disorder and an initial PHQ-9 score <10 were treated with usual primary care.

Cohort

We retrospectively reviewed the medical records of 4663 adult patients (aged 18–93 years) enrolled in the CCM from March 1, 2008, through June 30, 2015. Some patients (1567, or 33.6%) were missing 6-month follow-up PHQ-9 scores and were excluded from the analysis. Patients with complete data enrolled on or before December 31, 2013, served as the training set (n = 2525). Patients enrolled on or after January 1, 2014 till the end of the study, served as a validation set (n = 571), for a study cohort total of 3096 patients.

The dependent variable was the PHQ-9 score at the 6-month follow-up after enrollment in the CCM program. Remission was defined as a 6-month PHQ-9 score <5, whereas the presence of PDSs was defined as a PHQ-9 score ≥10. In addition to the initial diagnosis (first episode of major depression or recurrent major depression or dysthymia), the independent demographic variables were age, sex, and marital status (yes or no). Independent clinical variables included the results of the baseline PHQ-9 score, MDQ screen, Alcohol Use Disorders Identification Test (AUDIT) questionnaire,[23] and GAD-7 score.

The PHQ-9 is a 9-item tool that scores each item from 0 to 3 based on the patient's assessment over the prior 2-week period. The minimum score is 1 and the maximum is 27. The MDQ screen was scored as "negative" if the total score was <7 points for first question and if both the second and third questions had a negative response; it was scored as "abnormal" if any mix (or all) of the positive criteria were coded. As noted previously, patients with a clinical diagnosis of bipolar disorder were excluded from CCM enrollment, but no one was excluded based on a positive MDQ screen alone. The AUDIT has 10 questions that are scored from 0 to 4. An AUDIT score >8 signifies an increased risk of hazardous drinking.[24] The GAD-7 has a range of scores from 0 to 21, and anxiety symptomology was

defined as asymptomatic (score of 0–4), mild (5–9), moderate (10–14), or severe (≥15).

Classifiers

Using the training data set, 2 multivariate logistic regression classifiers were created: 1 for probability of remission and the other for the probability of PDSs. In both models, backward elimination of factors with a two sided P value of > .05 was used to arrive at a final simplified model in both cases. Analysis of variance with $P < .05$ for statistical significance was used to test the null hypothesis that the coefficients of the factors eliminated from the simplified model were not equal to zero.

Classifier Validation

Both classifiers obtained from the training data set using the simplified models were run on the validation data set to produce the predicted probabilities of remission from depression and PDSs. These predicted probabilities were compared with the actual 6-month PHQ-9 score to produce receiver operating characteristic (ROC) curves. The c-statistic, or area under the curve (AUC), was computed.[25] Bootstrapping using 1000 random samples with replacement of the validation data set was used to compute the 95% confidence intervals for the c-statistic in both classifiers. The statistical software R (version 3.0.2; R Foundation for Statistical Computing, Vienna, Austria), including the ROCR[26] and Boot[27] packages, were used to analyze the data.

Results

Of the 3096 patients in this study, 2525 (81.6%) were in the training cohort and the remainder were in the validation cohort (n = 571 [18.4%]). Remission rates at 6 months were 52.3% (n = 1321) for the training group and 49.7% (n = 284) for the validation group ($P = .266$). PDS was present in 594 patients (23.5%) in the training cohort and 152 patients (26.2%) in the validation cohort ($P = .123$). Mild depression severity (PHQ-9 score of 5–9) was present in 23.5% versus 26.6%, respectively.

Remission Classifier

The odds ratios and 95% confidence intervals for the factors retained in the final simplified remission of depressive symptoms classifier are illustrated in Figure 1. Sex, dysthymia diagnosis, and the AUDIT score were dropped from the initial model by backward selection because their P value is > .05. Comparing the final simplified model with the initial full model, the null hypothesis that coefficients of the dropped factors were not equal to zero was rejected ($P = .057$). Thus, the simplified model is not significantly different from the full model.

The remission of depressive symptoms classifier produced the ROC curve shown in Figure 2 when run using the validation data. The c-statistic (AUC) was 0.62, with a 95% confidence interval of 0.57 to 0.66.

Predicted probabilities for remission at 6 months using the validation data ranged from 14.5% to 79.1%, with a median of 50.6% (first quartile, 41.0%; third quartile, 59.3%).

Figure 1. Odds of a 9-item Patient Health Questionnaire (PHQ-9) score <5 at 6 months after enrollment in collaborative care management for depression, by variable. MDQ, Mood Disorder Questionnaire; GAD-7: 7-item Generalized Anxiety Disorder Questionnaire.

Variable	Odds Ratio (95% CI)
Age	1.01 (1.00 − 1.01)
Married	1.34 (1.14 − 1.58)
Recurrent Depression	0.70 (0.60 − 0.83)
Initial PHQ−9 Score	0.94 (0.92 − 0.96)
MDQ Negative	1.77 (1.43 − 2.19)
GAD−7 Score	0.97 (0.96 − 0.99)

Figure 2. Comparison of the Receiver Operating Characteristics (ROC) curves for Depression Outcome Validation using DOC-6 tool.

ROC Curves Predicting Six Month Outcome

Remission: AUC = 0.62, 95% CI: (0.57 – 0.66)
PDS: AUC = 0.67, 95% CI: (0.62 – 0.72)

Persistent Depressive Symptoms Classifier

Figure 3 shows the odds ratios and 95% confidence intervals for the factors retained in the final simplified persistent depressive symptoms classifier. Sex, dysthymia, and the AUDIT score were dropped from the initial model by backward selection because their *P* value was > .05. Comparing the final simplified model with the initial model, the null hypothesis that coefficients of the dropped factors were not equal to zero was rejected (*P* = .42). Thus, the simplified model is not significantly different from the full model.

The persistent depressive symptoms classifier produced the ROC curve shown in Figure 2 when

run using the validation data. The c-statistic (AUC) was 0.67, with a 95% confidence interval of 0.61 to 0.72. Predicted probabilities for PDSs on the validation data ranged from 5.5% to 73.1%, with a median of 23.5% (first quartile, 15.9%; third quartile, 32.9%).

Depression Outcomes Calculator

To ease the burden of referencing numerous coefficients and performing logarithmic calculations in a clinical setting, an online calculator was designed using the validated model. This tool was named the Depression Outcomes Calculator–Six Items (DOC-6) (© 2016 Angstman et al. All rights reserved) and can be accessed at http://www.

Figure 3. Odds of a 9-item Patient Health Questionnaire (PHQ-9) score ≥10 at 6 months after enrollment in collaborative care management for depression, by variable. MDQ, Mood Disorder Questionnaire; GAD-7: 7-item Generalized Anxiety Disorder Questionnaire.
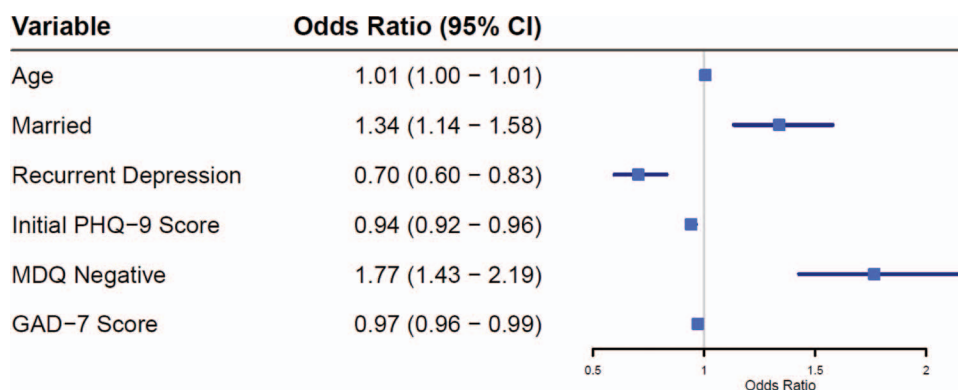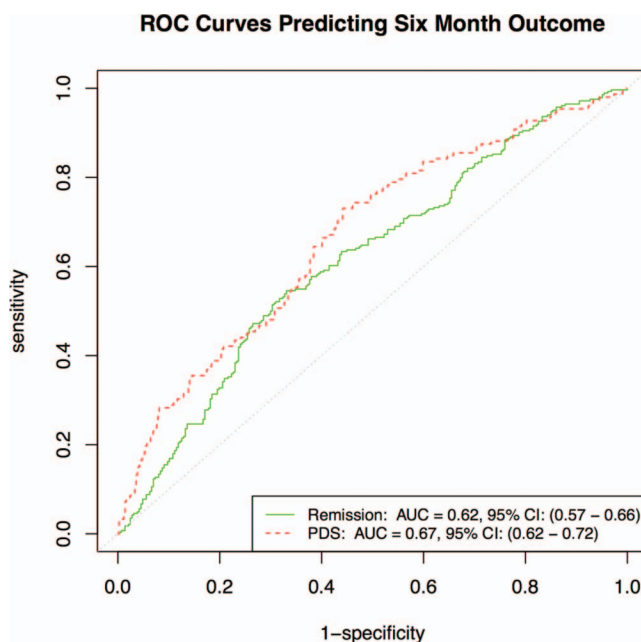
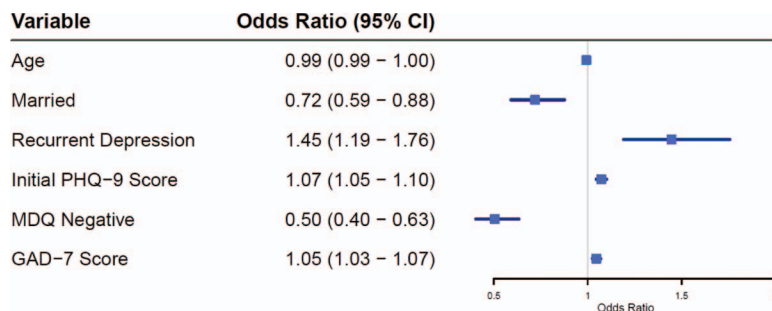| Variable	Odds Ratio (95% CI)
Age	0.99 (0.99 – 1.00)
Married	0.72 (0.59 – 0.88)
Recurrent Depression	1.45 (1.19 – 1.76)
Initial PHQ–9 Score	1.07 (1.05 – 1.10)
MDQ Negative	0.50 (0.40 – 0.63)
GAD–7 Score	1.05 (1.03 – 1.07)

Figure 4. Predictive probability of 6-month outcomes for a 20-year-old married patient with an uncomplicated (asymptomatic for anxiety symptoms and a negative Mood Disorder Questionnaire screen) first episode of depression, with an initial 9-item Patient Health Questionnaire score of 10.
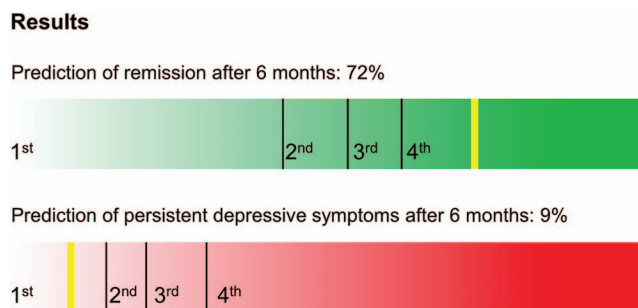
mayoclinic.org/doc-6. The online calculator was built using simple HTML5 form elements and the Knockout JavaScript framework (http://knockoutjs.com) to bind data between the data model and form fields. By prompting the user for each of the variable inputs through a simple form, the tool automatically applies the appropriate formulas and renders the output numerically and graphically, rounded to the nearest percentage. The use of standard web protocols and full client-side execution allows the maximum portability of this tool across browsers and devices. Results are given as the calculated predictive probability, with placement in the appropriate quartile. The graphical output is dynamically created upon execution and is accomplished using HTML styles in order to achieve instantaneous response times. The overall scale of the graphic display represents 0.0–1.0/300 pixels, which was chosen to optimize the ability to view it on mobile devices. The black quartile demarcations are a single pixel wide, making them accurate to the nearest third of a percentage. The yellow result line is 3 pixels wide, making it equivalent to 1% of the overall scale; the yellow line is placed on the scale to illustrate the percentage point.

Discussion

Using baseline characteristics and standardized assessment tools, we developed and preliminarily validated 2 logistic regression classifiers that could predict the probability of remission and PDSs at 6 months for depressed patients in a CCM, with moderate success (c-statistics of 0.62 and 0.67). We created a simple web-based calculation tool, the DOC-6, which uses these classifiers to deliver easily visualized probabilities and quartile comparisons. This prognostic tool allows care managers and physicians to triage patients based on a data-driven and validated risk stratification, alerting them to those patients who may require closer monitoring or additional therapeutic services.

Figures 4 and 5 demonstrate the probable predictive output of the model for 2 patient scenarios. The first is a 20-year-old married patient with an

Figure 5. Predictive probability of 6-month outcomes for a 67-year-old unmarried patient with a complicated (symptomatic for anxiety symptoms with a 7-item Generalized Anxiety Disorder Questionnaire score of 18 and a partially positive Mood Disorder Questionnaire screen) recurrent episode of depression and an initial 9-item Patient Health Questionnaire score of 26.
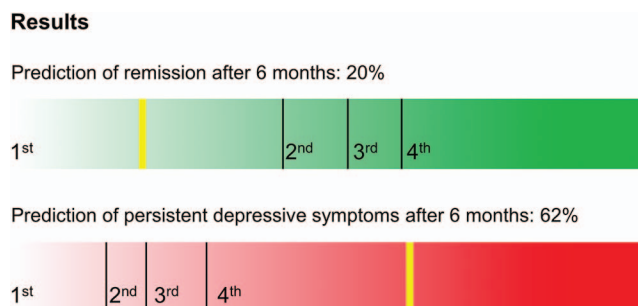
uncomplicated (asymptomatic for anxiety symptoms and a negative MDQ screen) first episode of depression, with an initial PHQ-9 score of 10 (Figure 4). The DOC-6 output demonstrates a 72% predicted probability of remission at 6 months (fourth quartile), with only a 9% probability of PDSs (first quartile). To contrast this example, Figure 5 illustrates the DOC-6 output for a 67-year-old unmarried patient with a complicated (symptomatic for anxiety symptoms, with a GAD-7 score of 18 and a partially positive MDQ screen) recurrent episode of depression and an initial PHQ-9 score of 26, with a 20% predicted probability of remission (first quartile) at 6 months and a 62% chance of PDSs (fourth quartile) at 6 months.

The DOC-6 is practical tool that uses results from standardized, widely available assessment tools (the PHQ-9, GAD-7, and MDQ). In our practice, all the relevant information was obtained from the patient by the care manager. Obtaining the data and inputting it into the tool would not require clinician activity. Thus, few barriers exist to implementation within primary care. We have demonstrated that other comorbid psychiatric conditions, such as posttraumatic stress disorder, can have a negative effect on depression outcomes in a CCM.[28] It is not always possible to make time available for a structured interview or even a full psychiatric history, especially in a busy primary care practice with limited psychiatric resources. Input for the DOC-6 can be obtained during a clinical visit or a phone call, or via other electronic methods. While the result does not predict the individual patient's exact outcome (as this is affected by other factors, especially current treatment availability), it will potentially allow the clinician the opportunity to more effectively stratify risk and triage those patients who may need more clinical or therapeutic support. Future research may wish to evaluate whether the use of the DOC-6 contributes to increased clinical outcomes and cost-effectiveness among the primary care patient population.

This study has several limitations. The patients were diagnosed with major depression or dysthymia by their primary care provider and reviewed by a consulting psychiatrist. A structured diagnostic interview such as the Structured Clinical Interview for DSM-5 was not used, but this is consistent with routine clinical care, making the results more generalizable. The DOC-6 was validated at a large academic institution using a cohort of community primary care patients. While we have no reason to believe that our primary care patient population varies significantly from other primary care settings across the United States, cross-validation in other geographic locations and within minority populations will contribute to the validity of the DOC-6. The patients included in this study were all in a care coordination program. Thus, the algorithm may not generalize exactly to depressed patients who are not receiving coordinated care, as CCM patients receive more intensive care than those in usual care. Predictive classification performance of the model was moderate (AUC = 0.62–0.67), but is similar to other complex clinical classification scenarios, such as predicting 30-day readmissions.[29,30] Inclusion of variables such as biomarkers or depression subtypes in future research may help refine the model, improving the predictive performance.

Conclusions

Using primary care patient data in a CCM for depression treatment, we were able to identify 6 readily available variables (2 demographic and 4 clinical), then validated them for the prediction of 6-month remission and PDSs. The DOC-6 tool may be used to predict which patients may be at risk for worse outcomes.

To see this article online, please go to: http://jabfm.org/content/30/3/281.full.

References

1. Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. Arch Intern Med 2006;166:2314–21.
2. Thota AB, Sipe TA, Byard GJ, et al. Collaborative care to improve the management of depressive disorders: a community guide systematic review and meta-analysis. Am J Prev Med 2012;42:525–38.
3. Unutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA 2002;288:2836–45.
4. Angstman KB, Shippee ND, Maclaughlin KL, et al. Patient self-assessment factors predictive of persistent depressive symptoms 6 months after enrollment in collaborative care management. Depress Anxiety 2013;30:143–8.
5. Hirschfeld RM, Calabrese JR, Weissman MM, et al. Screening for bipolar disorder in the community. J Clin Psychiatry 2003;64:53–9.

6. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999;282:1737–44.

7. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166:1092–7.

8. Angstman KB, Dejesus RS, Rohrer JE. Correlation between mental health co-morbidity screening scores and clinical response in collaborative care treatment for depression. Ment Health Fam Med 2010;7:129–33.

9. Angstman KB, Maclaughlin KL, Rasmussen NH, DeJesus RS, Katzelnick DJ. Age of depressed patient does not affect clinical outcome in collaborative care management. Postgrad Med 2011;123:123–45.

10. Angstman KB, MacLaughlin KL, Williams MD, Rasmussen NH, DeJesus RS. Increased anxiety and length of treatment associated with depressed patients who are readmitted to collaborative care. J Prim Care Community Health 2011;2:82–6.

11. Angstman KB, Pietruszewski P, Rasmussen NH, Wilkinson JM, Katzelnick DJ. Depression remission after six months of collaborative care management: role of initial severity of depression in outcome. Ment Health Fam Med 2012;9:99–106.

12. Gunn J, Elliott P, Densley K, et al. A trajectory-based approach to understand the factors associated with persistent depressive symptoms in primary care. J Affect Disord 2013;148:338–46.

13. Bosworth HB, Hays JC, George LK, Steffens DC. Psychosocial and clinical predictors of unipolar depression outcome in older adults. Int J Geriatr Psychiatry 2002;17:238–46.

14. Bosworth HB, McQuoid DR, George LK, Steffens DC. Time-to-remission from geriatric depression: psychosocial and clinical factors. Am J Geriatr Psychiatry 2002;10:551–9.

15. Rubenstein LV, Rayburn NR, Keeler EB, Ford DE, Rost KM, Sherbourne CD. Predicting outcomes of primary care patients with major depression: development of a depression prognosis index. Psychiatr Serv 2007;58:1049–56.

16. 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;163:1905–17.

17. Sherbourne C, Schoenbaum M, Wells KB, Croghan TW. Characteristics, treatment patterns, and outcomes of persistent depression despite treatment in primary care. Gen Hosp Psychiatry 2004;26:106–14.

18. Simon GE, Von Korff M, Lin E. Clinical and functional outcomes of depression treatment in patients with and without chronic medical illness. Psychol Med 2005;35:271–9.

19. King M, Weich S, Torres-González F, et al. Prediction of depression in European general practice attendees: the PREDICT study. BMC Public Health 2006;6:6.

20. Rifel J, Svab I, Ster MP, Pavlic DR, King M, Nazareth I. Impact of demographic factors on recognition of persons with depression and anxiety in primary care in Slovenia. BMC Psychiatry 2008;8:96.

21. Angstman KB, Rasmussen NH, Herman DC, Sobolik JJ. Depression care management: impact of implementation on health system costs. Health Care Manag (Frederick) 2011;30:156–60.

22. Williams M, Angstman K, Johnson I, Katzelnick D. Implementation of a care management model for depression at two primary care clinics. J Ambul Care Manage 2011;34:163–73.

23. Reinert DF, Allen JP. The alcohol use disorders identification test: an update of research findings. Alcohol Clin Exp Res 2007;31:185–99.

24. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption–II. Addiction 1993;88:791–804.

25. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. BMC Med Res Methodol 2012;12:82.

26. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005;21:3940–1.

27. Davison AC, Hinkley DV. Bootstrap methods and their application. New York: Cambridge University Press; 1997.

28. Angstman KB, Marcelin A, Gonzalez CA, Kaufman TK, Maxson JA, Williams MD. The impact of posttraumatic stress disorder on the 6-month outcomes in collaborative care management for depression. J Prim Care Community Health 2016;7:159–64.

29. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA 2011;306:1688–98.

30. Garrison GM, Robelia PM, Pecina JL, Dawson NL. Comparing performance of 30-day readmission risk classifiers among hospitalized primary care patients. J Eval Clin Pract 2016 Oct 3 [Epub ahead of print].