Predictors of diagnostic transition from major depressive disorder to bipolar disorder: a retrospective observational network study

Anastasiya Nestsiarovich, Jenna M. Reps, Michael E. Matheny, Scott L. DuVall, Kristine E. Lynch, Maura Beaton, Xinzhuo Jiang, Stephen R. Pfoh, Nigam H. Shah, Carmen Olga Torres, Christian G. Reich, Dong Yun Lee, Sang Joon Son, Seng Chan You, Rae Woong Park, Patrick B. Ryan and Christophe G. Lambert

INTRODUCTION

Mood disorders are one of the three leading causes of disability worldwide [1], with major depressive disorder (MDD) affecting more than 17 million Americans [2] and bipolar disorder (BD) affecting 7 million Americans annually [3]. Both MDD and BD are chronic debilitating psychiatric conditions, with overlapping neurobiology and symptoms (recurrent depressive episodes), however, BD is diagnosed if at least one manic/hypomanic episode was present during the patient lifetime [4]. It is debatable whether MDD is a comorbidity of BD or its earlier stage, and no consensus exists on individual conversion predictors, delaying BD’s timely recognition and treatment. We aimed to build a predictive model of MDD to BD conversion and to validate it across a multi-national network of patient databases using the standardization afforded by the Observational Medical Outcomes Partnership (OMOP) common data model. Five “training” US databases were retrospectively analyzed: IBM MarketScan CCAE, MDCR, MDCD, Optum EHR, and Optum Claims. Cyclops regularized logistic regression models were developed on one-year MDD-BD conversion with all standard covariates from the HADES PatientLevelPrediction package. Time-to-conversion Kaplan-Meier analysis was performed up to a decade after MDD, stratified by model-estimated risk. External validation of the final prediction model was performed across 9 patient record databases within the Observational Health Data Sciences and Informatics (OHDSI) network internationally. The model’s area under the curve (AUC) varied 0.633–0.745 (µ = 0.689) across the five US training databases. Nine variables predicted one-year MDD-BD transition. Factors that increased risk were: younger age, severe depression, psychosis, anxiety, substance misuse, self-harm thoughts/actions, and prior mental disorder. AUCs of the validation datasets ranged 0.570–0.785 (µ = 0.664). An assessment algorithm was built for MDD to BD conversion that allows distinguishing as much as 100-fold risk differences among patients and validates well across multiple international data sources.
onset, higher depression severity, multiple depressive episodes, family history of mood disorders, co-existing alcohol/substance abuse, attention-deficit/hyperactivity disorder, anxiety disorders, psychoses, suicide attempts, personality disorders, hospitalization, as well as rapid mood cycling, psychotherapy, living alone, prior use of psychotropic drugs, and others [8, 14–18]. Contradicting data are reported across the literature on the role of sex, age, depression onset, and severity in MDD-BD diagnostic transition.

This work set out to develop a predictive model of the conversion from MDD to BD, and to validate the model across a multi-national network of patient databases using the standardization afforded by the Observational Medical Outcomes Partnership (OMOP) common data model. The final goal was to develop a simple clinically useful risk assessment algorithm to help practitioners to recognize BD as early as possible among the patients presenting with MDD, and thus, to shorten the duration of untreated illness and mitigate the iatrogeny.

PATIENTS AND METHODS

This was a retrospective observational cohort study to develop a prognostic model of MDD diagnostic conversion to BD within a one-year period. Patient data on diagnoses, observations, provider visits, procedures performed, and medications filled were extracted from each data repository and analyzed within our computational health platform [19]. Five "training" US databases were used to develop the model: IBM MarketScan Commercial Claims and Encounters Database (CCAE, 2001–2018), IBM MarketScan Medicare Supplemental Database (MDCR, 2001–2018), IBM MarketScan Multi-State Medicaid Database (MDDC, 2005–2018), Optum de-identified Electronic Health Record Dataset (Optum EHR, 2005–2018), and Optum de-Identified Clinformatics Data Mart Database (Optum claims, 2001–2018). The sociodemographic and clinical data were extracted on all individuals who had a first observable diagnosis of MDD and then either had subsequent BD diagnosis within one year or not. Patient inclusion criteria were: age >10 years at the time of the first recorded MDD diagnosis ("index visit"); at least one year of observation before the index visit, no diagnosis of BD, schizophrenia, or schizoaffective disorder at any time point before the index visit (from the start of patient database coverage), and no antidepressant, antipsychotic, lithium, or mood-stabilizing anticonvulsant (MSA) use recorded in the given database at any time point before the index visit. A one-year observation period before the index visit was required to collect the relevant covariates on each patient to predict their future diagnostic transition. All selected individuals with MDD were followed up to 1 year, but if BD diagnosis occurred within this one-year period, the observation was stopped on the day that BD diagnosis was coded. Thus, patients with MDD-BD diagnosis conversion within one year were considered as "cases", and those who did not convert during one year after the index visit were considered as a "control group".

Cyclops regression models were developed using these covariates. We then retrained the model on each database using Cyclops with the composite covariates (see our electronic data source https://github.com/ohdsi-studies/BipolarMisclassificationValidation/tree/master/instr/cohorts for composite covariate definitions). For each composite covariate, an average coefficient was calculated across models for the five databases. These values were then multiplied by 10 and rounded to the nearest integer to enable an easily computable overall risk score per person by summing the rounded values across the final set of composite covariates.

The final regression model was then shared with our collaborators across the US and internationally for external validation. We also performed a sensitivity analysis on feature importance using our five training databases among new cases of MDD that appeared in 2019 and onwards (after model development), by examining the impact on AUC of including/excluding each variable in the model.

To explore how consistent the long-term risk of diagnostic transition was for each patient as a function of an estimated individual’s short-term (one year) risk, we performed secondary Kaplan–Meier survival analysis on the time from MDD to BD diagnosis conversion (with censoring) as a function of prediction model risk score, going forward into the future as far as 10 years from the index visit. Also, for each training and validation dataset, we generated a Cox regression model of time to MDD conversion (with right censoring) as a function of a smoothing spline of the per-person risk score, with a score of zero as reference. This way we could estimate how much the overall risk score was associated with risk of MDD conversion.

For each training dataset, Cyclops also built a “calibration plot” which uses locally estimated scatterplot smoothing (LOESS) to plot the actual/observed risk as a function of the predicted risk [20]. A model is considered well-calibrated when the predicted risks match the observed risks.

External validation of the final prediction model was performed using several validation datasets within the OHDSI network including Columbia University (CUMC), Ajou University in South Korea (AUSOM) [13], Stanford medicine Research data Repository (STARR) [21], IQVIA (including Germany, France, Belgium, and US records), Japan Medical Data Center database (JMDC) [11], as well as the US Veterans Health Administration EMR. This validation approach was similar to the one used in a recent OHDSI study [12]. The present study was approved by the institutional review boards (IRBs) of the respective collaborating institutions, where applicable. The Ajou University Hospital IRB number is AJIR-MED-MDB-20-034. Approval for the use of STARR is provided by the Stanford Institutional Review Board, protocol 53248. The use of IBM and Optum databases was reviewed by the New England IRB and was determined to be exempt from IRB approval.

The overall schema of the methods used is displayed in Fig. 1.

RESULTS

There were a total of 2,687,578 patients included in all five training databases meeting the study eligibility criteria. Patient characteristics are reported in Supplementary Table S1 as the same pharmacological agent. Cyclops regression was used which has a procedure employing regularized regression with cross-validation, which drops covariates that do not add to the model’s discriminative performance, and thus, reduces the pool of thousands of candidate predictors into several hundred variables. We also employed XGBoost modeling as an alternative to logistic regression to determine if better performance could be achieved by modeling interactions and nonlinear relationships. We described the model’s performance (i.e. its capacity to robustly discriminate between cases and controls) by reporting the receiver operator characteristic area under the curve (AUC) across five training datasets.

The second stage of the modeling process was data and human-driven. A domain expert dataset was re-examined to identify a reduced set of predictors and grouped them into “composite” covariates related to the same medical problem (value 1 was used to code “present” and 0 —“absent”). For example, a composite variable “substance misuse” would have a value of 1 if the patient had either alcohol abuse or opioid dependence in the year before the index visit. Individual components of such composite covariates were required to have the same direction and magnitude of effect across the five training datasets. Covariates with inconsistent directionality across the datasets, as well as those with a hazard ratio values close to one (thus, not clinically significant), were excluded from the analysis. Computational definitions for each “composite” covariate were then expanded to incorporate additional descendent terms using SNOMED and drug ingredient vocabulary hierarchies, and logistic regression models were built using these covariates. We then retrained the model on each database using Cyclops with the composite covariates (see our electronic data source https://github.com/ohdsi-studies/BipolarMisclassificationValidation/tree/master/instr/cohorts for composite covariate definitions). For each composite covariate, an average coefficient was calculated across models for the five databases. These values were then multiplied by 10 and rounded to the nearest integer to enable an easily computable overall risk score per person by summing the rounded values across the final set of composite covariates.

The final regression model was then shared with our collaborators across the US and internationally for external validation. We also performed a sensitivity analysis on feature importance using our five training databases among new cases of MDD that appeared in 2019 and onwards (after model development), by examining the impact on AUC of including/excluding each variable in the model.

To explore how consistent the long-term risk of diagnostic transition was for each patient as a function of an estimated individual’s short-term (one year) risk, we performed secondary Kaplan–Meier survival analysis on the time from MDD to BD diagnosis conversion (with censoring) as a function of prediction model risk score, going forward into the future as far as 10 years from the index visit. Also, for each training and validation dataset, we generated a Cox regression model of time to MDD conversion (with right censoring) as a function of a smoothing spline of the per-person risk score, with a score of zero as reference. This way we could estimate how much the overall risk score was associated with risk of MDD conversion.

For each training dataset, Cyclops also built a “calibration plot” which uses locally estimated scatterplot smoothing (LOESS) to plot the actual/observed risk as a function of the predicted risk [20]. A model is considered well-calibrated when the predicted risks match the observed risks.

External validation of the final prediction model was performed using several validation datasets within the OHDSI network including Columbia University (CUMC), Ajou University in South Korea (AUSOM) [13], Stanford medicine Research data Repository (STARR) [21], IQVIA (including Germany, France, Belgium, and US records), Japan Medical Data Center database (JMDC) [11], as well as the US Veterans Health Administration EMR. This validation approach was similar to the one used in a recent OHDSI study [12]. The present study was approved by the institutional review boards (IRBs) of the respective collaborating institutions, where applicable. The Ajou University Hospital IRB number is AJIR-MED-MDB-20-034. Approval for the use of STARR is provided by the Stanford Institutional Review Board, protocol 53248. The use of IBM and Optum databases was reviewed by the New England IRB and was determined to be exempt from IRB approval.

The overall schema of the methods used is displayed in Fig. 1.
average fractions of MDD to BD conversions for each of the model covariates per training database. Experimental employment of XGBoost versus Cyclops logistic regression did not result in higher AUC values, therefore all further models used Cyclops to develop parsimonious models. Also, no appreciable reduction in AUC was observed during the simplification of the model when moving from hundreds of covariates with Cyclops towards our final model. As a result of the human-modified/corrected data analysis, our final simple additive score model identified nine variables predicting one-year MDD transition to BD, where the age variable is represented by 11 groups. Each of the nine variables predicting one-year MDD transition to BD, where the age variable is represented by 11 groups. Each of the variables was assigned a score (positive score means a higher risk of diagnostic transition, and negative score means a lower risk of diagnostic transition): age at MDD onset 10–14 (+1); age 15–29 (+12); age 30–34 (+10); age 35–39 (+9); age 40–44 (+8); age 45–49 (+7); age 50–54 (+5); age 55–59 (+3); age 60–69 (+0, a reference group); age 70–74 (−3); age over 75 (−5); other mental disorders one year before but excluding the index visit (−5); severe MDD at index visit (−5); psychosis at MDD index visit (−10); other mental disorders one year before but excluding the index visit (−2); anxiety disorders and/or use of anti-anxiety drugs one year before and including the index MDD visit (−2); pregnancy one year before and during the index visit (−3); substance misuse one year before and during the index visit (−5); self-harm thoughts or actions one year before and during the index MDD visit (−9). Note that the covariate “patient sex” was dropped from the model at the final stage of analysis since it was not conferring sufficient risk for diagnosis transition (coefficient was close to zero).

Table 1 shows the prediction performance of the model on the five training datasets. The training model AUCs varied 0.633–0.745 with an average of 0.689 across the databases.

Figure 2 shows the >100-fold hazard ratio range for MDD-RD transition over the range of covariate risk scores in the CCAE database. Thus, in the CCAE database, for a patient with zero risk score, the hazard ratio of diagnostic transition would be one, and for a patient with a risk score of 25, it would be 15. Therefore, the latter patient would have a 15-times higher relative risk of transitioning from MDD to BD, compared to the former patient. Supplementary Figs. S1–S14 contain risk score-hazard ratio plots for other databases, which show a similar magnitude of effect for the overall risk score per patient. Supplementary Figs. S15–S19 contain training models’ calibration plots which show an overall monotonically increasing relationship between model-predicted risk and the observed transition fraction. The MDCR model underestimated the conversion risk since the smoothed actual proportion of conversions was consistently higher than predicted. Figure 3 shows the Kaplan–Meier survival curves broken out by risk score range for the CCAE database.

Supplementary Figs. S20–S28 contain graphs with Kaplan–Meier survival curves for all training and validation datasets (some collaboration sites did not provide their graphs to maximally protect patient confidentiality).

External validation of our final predictive model showed performance better than random prediction (AUC = 0.50) across all the validation databases (Table 2), with AUCs ranging 0.570–0.785 (average 0.664). Figure 4 demonstrates the model’s performance in different databases (training and validation) as a function of time. As a result of our integrative data analysis from different databases, we developed a simple, clinically meaningful algorithm to estimate the individual patient’s risk of MDD diagnosis transition to BD within one year after the index visit (Fig. 5). Interactive results are available for viewing within an R Shiny application at https://data.ohdsi.org/MDDinBipolar/.

Supplementary Table S2 shows the results of our covariate sensitivity analysis when the final prediction model was applied prospectively on our training databases for patients with MDD onset in 2019. The full model’s AUC and the AUC when each of the 9 variables was excluded are shown (age was excluded with all of its 5-year interval variables). Excluding a covariate typically resulted in lower AUC values or the same AUC values as in the full model, though only age was statistically significant within the larger Optum and CCAE databases.
DISCUSSION
In our final regression model, younger patient age, higher severity of the initial depressive episode, the presence of psychotic features during the index depression, and anxiety were predictive of MDD diagnostic transition to BD within one year.

The association of diagnostic transition with a younger age can have the following explanation: BD typically manifests at a
younger age, which is supported by previous studies aiming to distinguish MDD from BD [22, 23]. The association of mood episode severity with BD transition is consistent with other findings that the highest severity depressive episodes occur in BD-I, followed by BD-II and MDD [12]. This observation is also supported by results from the study of Goldberg et al., where bipolar converters were found to have more severe depressive symptoms or psychotic symptoms than non-converters [24], as well as by results of other prospective and retrospective studies [23–25]. Anxiety (as evidenced by an anxiety disorder diagnosis or the intake of anti-anxiety medication) was among the MDD-BD switching predictors in our study. A recent cross-sectional study has shown that the prevalence of comorbid anxiety disorders was significantly higher in patients with BD (53.2%) than in patients with MDD (37.2%) [26]. Another study found that comorbid anxiety in patients with mood disorders can serve as a marker of clinical severity [27].

Older age, mild baseline depression, and pregnancy were found to be predictors of a lower MDD-BD transition risk within one year in our study. The pregnancy association may be explained by (1) there is generally a high level of vigilance among healthcare providers regarding recognition of postpartum depression, thus, hypomania/mania symptoms can remain unrecognized in this category of patients, (2) pregnant/postpartum women might enroll disproportionately higher with obstetric services, than with psychiatric ones, even if they have already established mental healthcare, (3) pregnancy/delivery itself, pharmacological treatment of pregnancy-associated conditions, or pharmacotherapy of MDD with a selected set of non-teratogenic antidepressants during the perinatal period could hypothetically have a protective effect on mania/hypomania symptoms.

The clinical utility of our study results is in informing scientific discussions about nosological differences between BD and MDD and helping to recognize BD manifesting with depression relatively early. The latter remains a major diagnostic challenge in clinical psychiatry. Patient surveys found that one-third of BD respondents waited for more than ten years before receiving the correct diagnosis despite actively seeking help [28, 29], and saw an average of four physicians [29]. In its turn, lack of timely treatment of BD due to misdiagnosis can lead to an increased risk of suicide, long-lasting functional impairment, unemployment, family and legal issues, frequent hospitalizations, and higher healthcare costs than with appropriate BD treatment [30–33], and early switch to lithium or other mood-stabilizing agents in high-

Table 2. Performance of the final score-based model predicting diagnostic transition from major depressive disorder (MDD) to bipolar disorder (BD) within one year, in “validation” (US and international) datasets.

| Data source | AUC (95% CI) | N of patients who met inclusion criteria | N of patients with BD outcome in 1 year | Proportion of MDD patient population with BD outcome |
|-------------|-------------|----------------------------------------|---------------------------------|-----------------------------------------------|
| CUIMC (US EHR) | 0.570 (0.543–0.598) | 5611 | 457 | 8.145 |
| JMD (Japanese) | 0.610 (0.545–0.675) | 1303 | 67 | 5.142 |
| IQVIA DAFR (French EMR) | 0.615 (0.482–0.749) | 1910 | 17 | 0.890 |
| IQVIA DAFR (German EMR) | 0.628 (0.597–0.658) | 127,353 | 315 | 0.247 |
| STARR (US EHR) | 0.646 (0.613–0.678) | 27,266 | 290 | 1.064 |
| Veterans Health Administration (US EMR) | 0.670 (0.665–0.675) | 359,449 | 9246 | 2.570 |
| IQVIA Ambulatory (US EMR) | 0.691 (0.680–0.702) | 148,343 | 1548 | 1.044 |
| IQVIA Belgium (Belgian EMR) | 0.757 (0.602–0.912) | 667 | 7 | 1.050 |
| AUSOM (South Korean EMR) | 0.785 (0.724–0.847) | 2570 | 30 | 1.167 |

AUC area under the curve, N number, BD bipolar disorder, MDD major depressive disorder, EHR electronic health records, EMR electronic medical records, CUIMC Columbia University database (US), JMDJ Japan Medical Data Center database, DAFR IQVIA data from France, DAGER IQVIA data from Germany, STARR STAnford medicine Research data Repository (US), AUSOM Ajou University data from South Korea.

Fig. 4 Performance of predictive model for one-year MDD-BD diagnosis conversion depending on a data recording year. AUC—area under the curve. CCAE—IBM MarketScan Commercial Claims and Encounters Database (US), IQVIA_ambemr—IQVIA Ambulatory database for US, IQVIA_DAGER—IQVIA data for Germany, MDCD—IBM MarketScan Multi-State Medicaid Database, MDCR—IBM MarketScan Medicare Supplemental Database, Optum EHR—Optum de-identified electronic health record dataset, Optum claims—Optum De-Identified Clininformatics Data Mart Database, STARR—STAnford medicine Research data Repository, VA—US Veterans Administration database.

A. Nestsiarovich et al. Translational Psychiatry (2021) 11:642
In our risk assessment algorithm, we considered early onset of depression, presence of severe depression, and psychotic features as important risk factors. Among clinicians, there was a need to clarify questions about manic/hypomanic episodes.

Assigning each patient an overall “risk score” based on our proposed risk assessment algorithm might be a useful clinical tool. For example, to calculate the risk of one-year BD conversion in a 21-year-old person, you could use Fig.5 part 1 to calculate her score of 13 (12-5 + 1 + 5) and then use part 2 to convert the score into a predicted risk of ~4%

As shown in Fig.2, the risk of diagnostic conversion from MDD to BD can differ ~10-fold between a patient with zero risk score and a patient with a risk score of 20. The difference between the “worst” case scenario when a patient has all the “unfavorable” factors for transition (score = 144, corresponding to a 50% conversion rate) versus the “best” scenario when a patient has only “favorable” factors (score = -10 in a 75-year-old patient presenting with mild MDD and negligible probability of pregnancy at this age, corresponding to a < 0.5% conversion rate) gives us a > 100-fold range of risk (Fig. 5).

Our predictive model was successfully validated in several external datasets from different countries, which supports its potential applicability across different healthcare systems. However, its overall performance was modest (average 0.69 in training datasets, and 0.66 in validation datasets), and was influenced by individual database short characteristics and recording practices. The small sample sizes from some of the external validation sites (Belgium, France, Japan, South Korea) led to wide confidence intervals in the performance estimates. The simple score model performed worse in CUIMC, but this database appears to have a much higher outcome rate, so this may indicate a different type of patient population or perhaps some differences in data recording. The relatively better performance of the AUSOM database may be explained by chance and the relatively small population who fits the inclusion criteria. The German and French data had lower percentages of one-year diagnostic conversion, which may be due to the data being extracted from primary care datasets, not covering enough psychiatric services. The predictive model was quite stable over time (from year to year), except for the IQVIA DAGER database, which represents German data, and had a relatively small number of outcomes (315 spread over 17 years) that could have led to the wide variance in yearly risk estimates.

Our prospective covariate sensitivity analyses on the patients with MDD onset in 2019 from the training databases suggest that the model performance and covariate directionality are stable moving forward in time, though only the age covariate had a significant effect on model performance when one variable was considered at a time.

On average over all 14 databases, the one-year conversion rate from MDD to BD was 2.666%. While this number may appear relatively low, if it was to be sustained for 10 years, a conversion rate of 1-(1-0.0266)^10 = 23.6% might be expected in a decade. While our Kaplan–Meier curves suggest these rates tend to drop off after the first few years, we do observe that persons who fall on the high end of risk in our models can exceed a 1 in 3 chance of conversion within a decade, representing an important population to screen for BD diagnosis and treatment.

**Study limitations.** Our data were extracted from electronic health records and administrative claims data, which have known limitations including incomplete data recording, variations of diagnostic decision criteria used by providers and of granularity/amount of patient-reported information during each visit. The data were unavailable before the patient database enrollment or the database start date. The external validity of the model could be better overall, some validation datasets had comparable AUCs to those of the training sets, and some were too small to accurately assess. Because of the limitations of the health insurance claims data, factors such as laboratory test results were not included in this study. We acknowledge that there was the potential for overfitting in the development of the models, but this was mitigated by validation in external databases with comparable models’ performance. We also recognize that it was not necessarily the first episode of MDD/BD in a patient’s life captured in our data (some patients started observation at 65 years old). Since hypomanic symptoms are often not reported by a patient and are not accounted for when making a diagnosis, our database could miss a portion of BD type II cases.

**CONCLUSIONS.** Our approach produced a simple, clinically understandable model for predicting one-year risk of diagnosis conversion from MDD to BD.
BD that validates well across multiple international data sources. Early onset of MDD, presence of psychotic features, severe depression, substance misuse, and suicidality can serve as clinical predictors of prospective transition of a patient to BD diagnostic group and should get close attention from clinicians. Despite moderate AUC performance, our model can identify patients spanning a 100-fold magnitude difference in risk of MDD to BD transition. Accounting for interactions and nonlinear relationships between the variables using XGBoost did not result in higher AUC values, suggesting that substantial improvements in model performance are unlikely to be gained by more feature engineering of the training databases, though temporal relationships might be explored further. Because one can code and bill for MDD without specifying severity, we anticipate that models that go beyond claims records to incorporate physician notes and results of depression psychometric assessment may lead to more precise predictions of the risk of future conversion.

REFERENCES
1. GBD. Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392:1789-858. 2017
2. National Institute of Mental Health. NIMH » Major Depression. https://www.nimh.nih.gov/health/statistics/major-depression.shtml (accessed 11 Feb 2020).
3. National Alliance on Mental Illness. Mental Health By the Numbers. https://www.nami.org/learn-more/mental-health-by-the-numbers (accessed 11 Feb 2020).
4. WHO. World Health Organization. The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research. World Health Organization; 1993.
5. Fagioli A, Cuomo A. On the centrality of mixed features in mood disorders: listening to Kraepelin and Weygandt and moving forward. Bipolar Disord. 2017;19:704–5.
6. Angst J, Menkangas KR, Cui L, Van Meter A, Ajdacic-Gross V, Rössler W. Bipolar spectrum in major depressive disorders. Eur Arch Psychiatry Clin Neurosci. 2018;268:741–8.
7. Malhi GS, Bell E. Mixed states: an early warning sign of profound problems in our taxonomy? Acta Psychiatr Scand. 2019;139:489–92.
8. Muslner KL, Østergaard SD. Patterns and predictors of conversion to bipolar disorder in 91 387 individuals diagnosed with unipolar depression. Acta Psychiatr Scand. 2018;137:32-32.
9. Angst J, Sellaro R, Stassen HH, Gamma A. Diagnostic conversion from depression to bipolar disorders: results of a long-term prospective study of hospital admissions. J Affect Disord. 2005;84:149–57.
10. Kessing LV, Willer L, Andersen PK, Buhl JD. Rate and predictors of conversion from unipolar to bipolar disorder: A systematic review and meta-analysis. Bipolar Disord. 2017;19:324–35.
11. Li C-T, Bai Y-M, Huang Y-L, Chen Y-S, Chen T-J, Cheng J-Y, et al. Association between antidepressant resistance in unipolar depression and subsequent bipolar disorder: cohort study. Br J Psychiatry. 2012;200:45–51.
12. Moreno C, Hasin DS, Arango C, Oquendo MA, Vieta E, Liu S, et al. Depression in bipolar disorder versus major depressive disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Bipolar Disord. 2012;14:271–82.
13. Truman CJ, Goldberg JF, Ghaemi SN, Baldassano CF, Wisniewski SR, Dennehy EB, et al. Self-reported history of manic/hypomanic switch associated with antidepressant use: data from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). J Clin Psychiatry. 2007;68:1472–9.
14. Hu Y-H, Chen K, Chang J-C, Shen CC. Critical predictors for the early detection of conversion from unipolar major depressive disorder to bipolar disorder: nationwide population-based retrospective cohort study. JMR Med Inf. 2020;8:e14278.
15. Frankland A, Roberts G, Holmes-Preston E, Perich T, Levy F, Lenroot R, et al. Clinical predictors of conversion to bipolar disorder in a prospective longitudinal familial high-risk sample: focus on depressive features. Psychol Med. 2018;48:713–21.
16. Kim H, Lim Y, Baek JH, Fava M, Mischoulon D, Nierenberg AA, et al. Predictive factors of diagnostic conversion from major depressive disorder to bipolar disorder in young adults ages 19-34: A nationwide population study in South Korea. J Affect Disord. 2020;265:52-58.

ACKNOWLEDGEMENTS
Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under award number R56MH120826. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

AUTHOR CONTRIBUTIONS
A.N., JMR, P.S.R., and C.G.L. designed the study. C.G.L. provided supervision. J.M.R. created the network study source code, source repository, and R Shiny application. A.N., J.M.R., and C.G.L. conducted the analyses, drafted the manuscript, and revised the manuscript. M.E.M., S.L.D., K.E.L., M.B., X.J., M.S., S.R.P., N.H.S., C.O.T., C.G.R., D.Y.L., S.J.S., S.C.Y., R.W.P., and P.B. contributed to model validation using their respective databases. All authors critically reviewed the manuscript, approved drafts, and approved the final paper.
COMPETING INTERESTS
J.M.R. and P.B.R. are employees of Janssen Research & Development and shareholders of Johnson & Johnson. C.G.R. and C.O.T. are employees of IQVIA. All other authors declare no conflict of interest.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01760-6.

Correspondence and requests for materials should be addressed to Christophe G. Lambert.

Reprints and permission information is available at http://www.nature.com/reprints

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

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

© The Author(s) 2021