ORIGINAL RESEARCH

# Derivation and validation of a multivariable model to predict when primary care physicians prescribe antidepressants for indications other than depression

Jenna Wong
Michal Abrahamowicz
David L Buckeridge
Robyn Tamblyn

Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada

**Objective:** Physicians commonly prescribe antidepressants for indications other than depression that are not evidence-based and need further evaluation. However, lack of routinely documented treatment indications for medications in administrative and medical databases creates a major barrier to evaluating antidepressant use for indications besides depression. Thus, the aim of this study was to derive a model to predict when primary care physicians prescribe antidepressants for indications other than depression and to identify important determinants of this prescribing practice.

**Methods:** Prediction study using antidepressant prescriptions from January 2003–December 2012 in an indication-based electronic prescribing system in Quebec, Canada. Patients were linked to demographic files, medical billings data, and hospital discharge summary data to create over 370 candidate predictors. The final prediction model was derived on a random 75% sample of the data using 3-fold cross-validation integrated within a score-based forward stepwise selection procedure. The performance of the final model was assessed in the remaining 25% of the data.

**Results:** Among 73,576 antidepressant prescriptions, 32,405 (44.0%) were written for indications other than depression. Among 40 predictors in the final model, the most important covariates included the molecule name, the patient's education level, the physician's workload, the prescribed dose, and diagnostic codes for plausible indications recorded in the past year. The final model had good discrimination (concordance ($c$) statistic 0.815; 95% CI, 0.787–0.847) and good calibration (ratio of observed to expected events 0.986; 95% CI, 0.842–1.136).

**Conclusion:** In the absence of documented treatment indications, researchers may be able to use health services data to accurately predict when primary care physicians prescribe antidepressants for indications other than depression. Our prediction model represents a valuable tool for enabling researchers to differentiate between antidepressant use for depression versus other indications, thus addressing a major barrier to performing pharmacovigilance research on antidepressants.

**Keywords:** antidepressant; indications; predictive studies; predictors; primary care; pharmacovigilance

Correspondence: Jenna Wong
Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 401 Park Drive, Suite 401 East, Boston, MA 02215, USA
Tel +1 617 867 4513
Fax +1 617 867 4276
Email jenna_wong@harvardpilgrim.org

## Introduction

Adverse drug reactions are a major cause of mortality, morbidity, and hospitalizations in the USA[1–3] and Canada.[4] Pharmacovigilance – defined by The World Health Organization as "the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems"[5] – is essential for increasing drug safety and reducing the frequency of adverse drug events.

**457**

An important component of pharmacovigilance is monitoring the use of medicinal products for unapproved indications.[6] Antidepressants are a group of drugs in need of heightened pharmacovigilance because they are one of the most commonly used medications[7–9] that are frequently prescribed for indications other than depression, many of which are unapproved or "off-label" for the drug.[10] Moreover, most off-label uses for antidepressants are not backed by a strong level of scientific evidence,[11] raising concerns because unsubstantiated off-label drug use increases the risk of adverse drug events and places unnecessary financial burdens on the health care system.[12] The frequency with which antidepressants are prescribed for non-evidenced-based indications highlights the need to conduct post-market evaluations of antidepressant-related outcomes by indication.

Administrative health databases have the potential to be valuable resources for studying antidepressant use for different indications. Such databases can identify large, population-based cohorts of antidepressant users, capture many different uses (ie, drug-indication combinations), and detect rare adverse drug events – all at a fraction of the cost required to conduct clinical trials, which tend to include smaller, more selective patient populations.[13] However, a major barrier precluding the use of administrative databases for this purpose is that they lack information on treatment indications, which is essential for distinguishing between antidepressant use for depression versus other indications. In the absence of documented treatment indications, studies[14–18] have inferred the reasons for antidepressant use based on the presence of diagnostic codes for plausible indications in administrative databases. We previously evaluated the accuracy of this approach and found that when compared to physician-documented treatment indications from an electronic prescribing system, administrative diagnostic codes had poor sensitivity (ie, identified only a small proportion of antidepressant prescriptions truly written for the corresponding indication). For example, diagnostic codes for depression, anxiety/stress disorders, and sleeping disorders identified only 26.5%, 31.2%, and 4.9% of antidepressant prescriptions for these indications, respectively.[19]

Unless other variables in health services data besides diagnostic codes can improve the ability to predict antidepressant treatment indications, such data will likely remain an unviable resource for studying antidepressant use by indication. Thus, the primary objective of this study was to derive and validate a model using a wide range of variables derived from health services data to predict when primary care physicians prescribed antidepressants for indications other than depression and to compare the performance of this model to that of a model based on diagnostic codes alone. A second and equally important objective of this study was to identify important predictors of antidepressant prescriptions for indications other than depression and measure their association with the outcome.

# Methods

## Study design and context

This prediction study took place in the Canadian province of Quebec, where a universal health insurance program covers all residents for the cost of medically necessary services, including physician visits, hospitalizations, diagnostic and therapeutic procedures, and psychiatric treatments.[20] Over 90% of physicians are reimbursed on a fee-for-service basis, with physicians submitting claims to the provincial health insurance agency (the Régie de l'assurance maladie du Québec [RAMQ]) for services rendered in hospitals or private clinics.[21] Each medical claim contains a billing code for the service rendered and an optional International Classification of Diseases, Ninth Revision (ICD-9) code for the primary complaint. The province also maintains a hospitalization discharge summary database (MED-ECHO) containing details of all hospitalizations at acute care institutions in Quebec, including in-hospital procedures and discharge diagnoses recorded by medical archivists based on structured chart abstraction. Before April 2006, discharge diagnoses in hospital abstracts were recorded using the ICD-9 system and in-hospital procedures were recorded using the Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP) system. Since April 2006, the ICD-10 coding system and the Canadian Classification of Health Interventions (CCI) system have been used.

By law, all Quebec residents must be covered for prescription drugs through either private plans (ie, group or employee benefit plans) or the public drug insurance plan. Approximately 50% of residents are registered in the public drug insurance plan, including individuals aged 65 years or older, welfare recipients, and those not insured through an employer.

## The Medical Office of the XXIst Century (MOXXI)

The Medical Office of the XXIst Century (MOXXI) is an indication-based electronic prescribing and drug management system used by consenting primary care physicians at community-based clinics around two major urban centers in Quebec.[22] For all prescriptions, MOXXI physicians are

required to document at least one treatment indication using either a drop-down menu containing on-label and off-label indications (without distinction) or by typing the indication(s) into a free-text field. These physician-documented indications were previously validated against a blinded, post-hoc physician-facilitated chart review, and were shown to have excellent sensitivity (98.5%) and high positive predictive value (97.0%).[23]

Since 2003, 207 physicians (25% of eligible) and over 100,000 patients (26% of all who visited a MOXXI physician) have consented to participate in MOXXI and have their information used for research purposes. MOXXI physicians tend to be younger and have lower patient loads than non-MOXXI physicians, while MOXXI patients tend to be older with more health complexities than non-MOXXI patients.[24]

## Data sources

This study included all prescriptions of drugs approved for depression that were written by a MOXXI physician between 1 January 2003 and 31 December 2012. The unit of analysis was the prescription. Patients with eligible prescriptions were linked by unique patient identifiers to beneficiary and medical claims data from RAMQ and hospital discharge abstracts in the MED-ECHO database. All databases accessed contained only anonymous or de-identified data. This study was approved by the McGill Institutional Review Board.

## Exclusion criteria

Drugs were excluded if they had fewer than 120 prescriptions in the MOXXI system over the study period (roughly corresponding to a prescribing frequency of less than once per month). As a result, all prescriptions for moclobemide, maprotiline, phenelzine, nefazodone, and tranylcypromine were excluded from the analysis.

## Study outcomes

The primary outcome was a binary variable representing whether the physician had prescribed the antidepressant for an indication other than depression. The secondary outcome was a polytomous variable that assigned antidepressant prescriptions to one of five treatment indication classes: 1) depression; 2) anxiety/stress disorders; 3) sleeping disorders; 4) pain; or 5) miscellaneous indications. Both outcomes were determined using the physician-documented indications in the MOXXI system, where treatment indications were grouped into categories using definitions under the ICD-9 system. For prescriptions where physicians recorded more

than one indication (1.5% of all antidepressant prescriptions), we used the indication entered first since it most likely represented the main or most responsible indication.

## Candidate predictors

We considered a wide range of characteristics related to the prescription, patient, or prescribing physician. Table 1 shows all 373 variables that were considered as candidate predictors of antidepressant prescriptions for indications other than depression.

## Prescription factors

We used information from the index prescription to create variables for the name of the prescribed molecule, the prescribed dose (in mg/day), whether the drug was to be taken on an "as-needed" basis, and the number of other drugs concurrently prescribed with the antidepressant. We considered these factors because certain antidepressants like selective serotonin reuptake inhibitors are more commonly prescribed for depression[10,25] but may be taken "as-needed" when used for other conditions like sexual dysfunction.[26] We also hypothesized that the prescribed dose was an important predictor because drugs like trazodone, amitriptyline, and doxepin are typically taken at lower-than-usual doses when prescribed for insomnia.[27]

## Patient factors

Patient factors included variables related to demographics, socio-economic status (SES), health problems (diagnostic codes), health services use, and drugs prescribed in the past year. SES was measured using area-level markers of income (median household income), education (% of population aged 25 to 64 years with less than university education), and employment status (% of population aged 25+ years without employment), which were determined by mapping the first three digits of patients' postal codes to their associated census tract divisions and calculating a weighted average of the respective estimates from the 2006 Canadian census. Patients' type of drug insurance (public or private plan) was also considered since residents under 65 years old with public drug insurance are typically unemployed or on welfare.

Diagnostic codes from physician billings or hospital discharge abstracts in the past year were expressed using 129 binary variables. Twenty-six of these variables captured the presence of diagnostic codes for 13 plausible treatment indications for antidepressants: depressive disorders, anxiety/stress disorders, pain, migraine, fibromyalgia, obsessive-compulsive disorder, vasomotor symptoms of menopause

**Table 1** Candidate predictors of antidepressant prescriptions for indications besides depression

| Variable | Values or FP1 function[a] |
|---|---|
| **Prescription-related factors (n=4)** | |
| Molecule name | 19 levels[b] |
| Prescribed dose (mg/day) | $X^1$ |
| Drug prescribed on a "take-as-needed" basis | Yes vs no |
| No. other drugs concurrently prescribed with the index drug | $X^{-2}$ |
| **Patient-related factors (n=362)** | |
| **Demographics and socio-economic status** | |
| Sex | Male vs female |
| Age (years) | $X^{-2}$ |
| Household income[c] (CAD) | $X^1$ |
| Less than university education[d] (%) | $X^1$ |
| Unemployment rate[e] (%) | $X^1$ |
| Type of drug insurance | Public vs private plan |
| **Diagnostic codes in the past year** | |
| Plausible antidepressant treatment indications[f] | |
| ±3 days around the index prescription date | 13 binary variables |
| 4 to 365 days before the index prescription date | 13 binary variables |
| Chronic conditions in the Charlson comorbidity index[g] | 17 binary variables |
| Other morbidities[h] | 86 binary variables |
| **Health services use in the past year** | |
| Number of outpatient visits | $X^{-0.5}$ |
| Number of outpatient physicians seen | $X^{-2}$ |
| Continuity of care with the prescribing physician[i] (%) | $X^1$ |
| Previous hospitalization | Yes vs no |
| Previous day surgery | Yes vs no |
| Previous ER visit | Yes vs no |
| Medical services[j] | 52 binary variables |
| In-hospital procedures[k] | 70 binary variables |
| **Drugs prescribed in the past year[l]** | 99 binary variables |
| **Physician-related factors (n=7)** | |
| Sex | Male vs female |
| Place of medical training | Canada/US vs other |
| Experience (years in practice) | 3 levels[m] |
| Workload (average no. patients per working day) | $X^{-0.5}$ |
| Factors affecting physician response to new information on evidence-based clinical practice[n] | |
| Evidence score | $X^3$ |
| Nonconformity score | $X^{-2}$ |
| Practicality score | $X^1$ |

**Notes:** [a]FP1 functions ($X^p$) are shown for continuous variables. For each continuous variable $X$, we selected the best fitting FP1 function among eight candidate FP1 functions where the powers $p$ were represented by the set {−2, −1, −0.5, 0, 0.5, 1, 2, 3} and $X^0$ denoted log($X$). In cases where the best $p \le 0$ and the variable's domain included 0, the original values of the variable were shifted up by 1 before applying the power. [b]Prescriptions were assigned to one of 19 levels: venlafaxine, duloxetine, desvenlafaxine, citalopram, paroxetine, escitalopram, sertraline, fluoxetine, fluvoxamine, amitriptyline, doxepin, nortriptyline, trimipramine, imipramine, desipramine, clomipramine, trazodone, bupropion, or mirtazapine. [c]Area-level measure representing the median household income (CAD) in the patient's census tract area. [d]Area-level measure representing the percentage of adults in the patient's census tract area with less than university education. [e]Area-level measure representing the percentage of unemployed adults in the patient's census tract area. [f]For each observation window, 13 binary variables were used to represent whether diagnostic codes were recorded for each of the following treatment indication categories: depression, anxiety/stress disorders, sleeping disorders, pain, migraine, fibromyalgia, obsessive-compulsive disorder, vasomotor symptoms of menopause, nicotine dependence, attention deficit/hyperactivity disorder, sexual dysfunction, pre-menstrual dysphoric disorder, and eating disorders. ICD-9 codes for these treatment indications are listed in Table S1. [g]17 binary variables were used to represent whether diagnostic codes for any of the following conditions in the Charlson comorbidity index were recorded in the past year: myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatic disease, peptic ulcer disease, mild liver disease, diabetes without chronic complication, diabetes with chronic complication, hemiplegia or paraplegia, renal disease, any malignancy, moderate or severe liver disease, metastatic solid tumor, and AIDS/HIV. ICD codes for these conditions were identified using the algorithms published by Quan et al.[31] [h]86 binary variables were used to represent each four-digit ICD-9 code that was recorded for at least 1% of all antidepressant prescriptions in the past year (after excluding diagnostic codes for antidepressant treatment indications and Charlson conditions). [i]Expressed as the percentage of all outpatient visits in the past year that were made to the prescribing physician. [j]Based on billing codes recorded in medical claims data over the past year. Individual billing codes were grouped into broader "billing code categories" using mapping tables obtained from the RAMQ. Binary variables were used to represent the presence of billing codes from any category that was recorded for at least 1% of antidepressant prescriptions in the past year (a total of 52 categories). [k]Based on procedure codes recorded in hospital discharge abstracts over the past year. Binary variables were used to represent the presence of any three-digit CCP code that was recorded for at least 1% of antidepressant prescriptions where the patient had been hospitalized in the past year (a total of 70 procedure codes). [l]Binary variables were used to represent the presence of a prescription in the past year for any drug (generic name) that had been prescribed in the past year for at least 1% of all antidepressant prescriptions (a total of 99 drugs). [m]Prescriptions were assigned to one of three levels: 1) 24+ years, 2) 15–23 years, or 3) <15 years. [n]Measured using physician scores on the Evidence-Nonconformity-Practicality survey,[33] which is a psychometric instrument for determining how physicians would likely respond to new information about evidence-based clinical practice. Higher evidence scores indicate a stronger belief in scientific evidence over clinical experience as the best source of clinical knowledge, higher nonconformity scores indicate more willingness to diverge from group norms in clinical practice, and higher practicality scores indicate higher sensitivity to practical concerns such as managing workload and patient flow. **Abbreviation:** FP1, first-degree fractional polynomial.

(ie, hot flashes), nicotine dependence, attention deficit/hyperactivity disorder, sexual dysfunction, pre-menstrual dysphoric disorder, and eating disorders. We examined two separate observation windows for these 13 diagnostic categories: 1) ±3 days around the index prescription date, and 2) 4 to 365 days before the index prescription date. We used the set of ICD-9 codes from our previous study[19] to identify diagnostic codes that mapped to each of these 13 diagnostic categories (Table S1). ICD-10 codes recorded in hospital discharge abstracts from April 2006 onwards were translated to their ICD-9 equivalent using conversion tables.[28] We also created another 17 binary variables to capture the presence of diagnostic codes for chronic conditions in the Charlson comorbidity index recorded over the past year since patients with depression often have more chronic morbidities than patients without depression.[29,30] Diagnostic codes for these chronic conditions were identified using the ICD mappings by Quan et al.[31] The last 86 binary variables captured the presence of all remaining four-digit ICD-9 codes that were recorded for at least 1% of all patients with antidepressant prescriptions in the past year (Table S2).

Patterns of health care utilization in the past year were also considered because compared to individuals without depression, individuals with major depression who are treated in primary care are more likely to have a usual source of care, more likely to see a physician, and less likely to visit the emergency room for routine services.[30] We therefore created variables representing the number of outpatient visits, number of outpatient physicians seen, and whether the patient had been hospitalized, undergone day surgery, or visited the ER in the past year. Continuity of care with the prescribing physician was also measured by calculating the percentage of all outpatient visits in the past year that were made to the prescriber. Medical services received in the past year were represented by assigning billing codes from physician claims data to their respective "billing code category" using mapping tables obtained from RAMQ. Binary variables were used to represent the presence of billing codes from any category recorded for at least 1% of patients with antidepressant prescriptions in the past year (a total of 52 categories – Table S3). A similar method was used for in-hospital procedures, where procedure codes from discharge abstracts in the past year were grouped at the three-digit CCP level. CCI codes in discharge abstracts after March 2006 were translated to their CCP equivalent using conversion tables from RAMQ. Binary variables were used to represent the presence of any 3-digit procedure code recorded for at least 1% of patients with antidepressant

prescriptions who had been hospitalized in the past year (a total of 70 procedure codes – Table S4).

Finally, we considered drugs prescribed to patients in the past year because patients' previous drug history may contain clues about the reasons for antidepressant prescriptions. For example, antidepressants may be more likely to be prescribed for depression if the patient was previously prescribed atypical antipsychotics (eg, aripiprazole, quetiapine) or lithium, which are drugs commonly used to augment antidepressant therapy for depression.[32] Binary variables were used to represent the presence of previous prescriptions for any molecule that had been prescribed in the past year for at least 1% of patients with antidepressant prescriptions (a total of 99 drugs – Table S5).

## Physician factors

Various factors may influence why physicians prescribe antidepressants for different indications, such as their previous medical training, clinical experience, workload, and receptiveness to new information on best practices. To test these hypotheses, we used information from RAMQ to capture physicians' sex, place of medical training (Canada/USA or other), level of clinical experience (number of years in practice since medical graduation), and workload (average number of patients seen per working day in the previous year). MOXXI physicians also completed the Evidence-Nonconformity-Practicality survey,[33] which is a psychometric instrument that determines how physicians will likely respond to new information about evidence-based clinical practice. Higher evidence scores (possible range 6–30) indicate a stronger belief in scientific evidence versus clinical experience as the best source of clinical knowledge, higher nonconformity scores (possible range 6–30) indicate more willingness to diverge from group norms in clinical practice, and higher practicality scores (possible range 5 to 25) indicate higher sensitivity to practical concerns such as managing workload and patient flow.

## Statistical analysis

Figure 1 illustrates the steps of the study analysis. For the primary outcome, we used multivariable binomial logistic regression to model the probability that a given prescription was written for an indication other than depression. For the secondary outcome, we used multivariable multinomial logistic regression to estimate the probabilities that a given prescription was written for each of the five mutually exclusive treatment indication classes.

Only a small proportion (5.1%) of prescriptions had missing data, which we excluded from the main analysis

(Figure 1). Missing data occurred either because the patient's postal code did not fall within a census tract region or because the prescribing physician did not complete the Evidence-Nonconformity-Practicality survey. We hypothesize that the mechanisms behind missing data were not related to factors affecting antidepressant treatment indications.

All prescriptions with complete data were randomly divided using a 3:1 split into a "training set" versus "test set". The training set was used for model selection and to fit the final prediction models. The test set was only used to evaluate the performance of the final models; it was not used in the model building or estimation process. Because prescriptions were nested within patients who were in turn nested within physicians, we assigned a random sample of 75% of physicians (rather than individual prescriptions) to the training set. The remaining 25% of physicians were assigned to the test set. All prescriptions from the same physician and for the same patient were therefore limited

to either the training or test set. To ensure that patients and prescriptions were also divided approximately 3:1 between the training and test sets, we first stratified physicians by the number of their patients and then randomly sampled separately within each stratum.

Using the same randomization procedure, we divided the physicians in the training set into three mutually exclusive blocks (Figure 1, Step 1). These blocks were used for 3-fold cross-validation to reduce the risk of overfitting the final model to the training set.[34] The cross-validation procedure involved fitting a candidate model for the primary outcome using data from two of the three blocks (the "derivation set") and evaluating its performance in the held-out block (the "validation set") (Figure 1, Step 2 – Box). We repeated this process three times, each time using a different block as the validation set and then averaged the performance over the three validation sets. As the performance metric, we used the scaled Brier score[35,36], similar to the $R^2$ statistic in
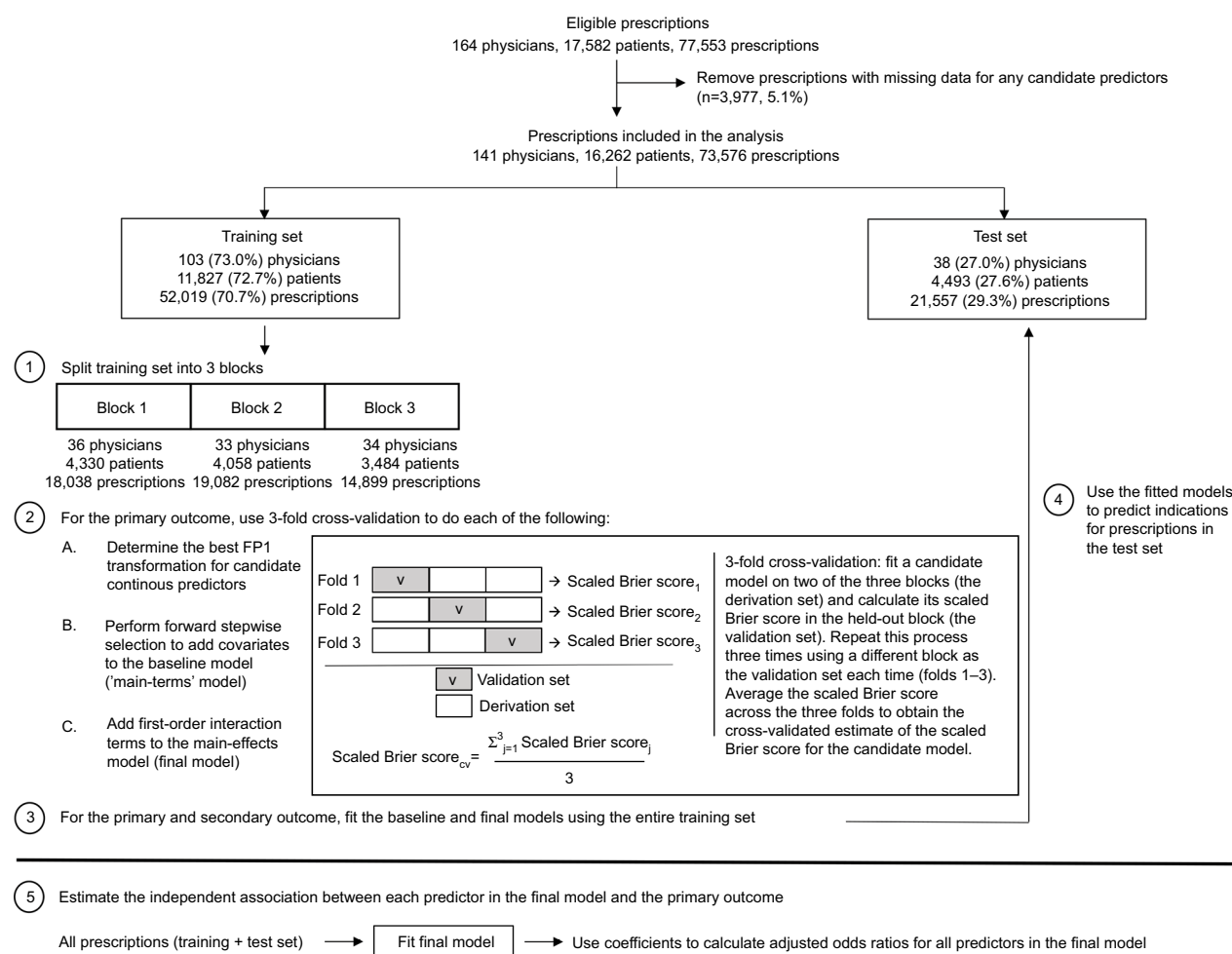


**Figure 1** Outline of the study analysis.

linear regression, which was calculated using the following formula:

$$\text{Brier score}_{\text{scaled}} = 1 - \left( \frac{1}{N} \sum_{i=1}^{N} \left( \hat{Y}_i - Y_i \right)^2 \bigg/ \frac{1}{N} \sum_{i=1}^{N} \left( \bar{Y} - Y_i \right)^2 \right)$$

where $N$ represents the total number of antidepressant prescriptions, $\hat{Y}_i$ represents the predicted probability that prescription $i$ was written for an indication other than depression, $Y_i$ represents the observed outcome for prescription $i$ (1 if the prescription was not written for depression, 0 otherwise), and $\bar{Y}$ represents the overall (marginal) observed probability of an antidepressant prescription being written for an indication other than depression in the study sample. As the formula shows, the scaled Brier score can be interpreted as the proportion by which a given model reduces the mean squared error of a non-informative model where all prescriptions are assigned the overall probability of having an indication other than depression.

We used 3-fold cross-validation to guide our decisions in three aspects of the model building process. The first aspect concerned the choice of functional form for the association between each continuous variable and the primary outcome (Figure 1, Step 2A). To this end, we relied on the flexible yet parsimonious first-degree fractional polynomials (FP1)[37]. For each continuous variable $X$, we selected the best fitting FP1 function among eight candidate FP1 functions: $X^p$, where the powers $p$ were represented by the set $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$, and $X^0$ denoted $\log(X)$[37]. The best functional form for each $X$ was assumed to be $X^1$ (ie, a linear association with outcome) unless a model using one of the nonlinear FP1 functions had a cross-validated scaled Brier score that was at least 0.0005 units higher than a model using the linear function. We required this minimum increase in the scaled Brier score to avoid using more complex functional forms that offered only minimal improvement in performance. The second decision concerned the selection of covariates. Starting with a model containing only covariates based on diagnostic codes for plausible antidepressant treatment indications (the "baseline model"), we used a forward stepwise selection procedure to add other covariates to the model (Figure 1, Step 2B). At each step, the variable that produced the greatest increase in the cross-validated scaled Brier score was added to the model. We stopped adding covariates when none of the remaining variables further increased the cross-validated scaled Brier score by at least 0.0005 (again, to avoid including covariates that offered only minimal improvement in predictive performance). Finally, among the covariates added to the

model from the forward stepwise selection procedure, we identified plausible first-order interaction terms between them and tested whether the individual addition of these interaction terms improved the cross-validated scaled Brier score of the "main-terms" model (ie, without interactions) by at least 0.0005 (Figure 1, Step 2C).

After identifying the final prediction model, we used the entire training set to fit both the baseline and final models for the primary and secondary outcomes (Figure 1, Step 3). We then used the fitted coefficients of these models to calculate the probability of the outcome for prescriptions in the test set, where the outcome probability was calculated as $1 / \left( 1 + e^{-(intercept + \beta_1 X_1 + \dots + \beta_k X_k)} \right)$. We evaluated the performance of these predictions using the methods and criteria described below (Figure 1, Step 4).

## Measures of predictive performance
### Primary outcome
We assessed overall model performance using the scaled Brier score. We assessed model discrimination (the model's ability to distinguish between prescriptions for depression versus other indications) using two measures: 1) the concordance ($c$) statistic, which is equivalent to the area under the receiver operating characteristic curve; and 2) the discrimination slope, which is calculated as the absolute difference in the average predicted probabilities among prescriptions for depression and prescriptions for other indications.[36] We compared the discrimination of the final and baseline models using the integrated discrimination improvement (IDI) to quantify the predictive ability gained from including other health-related information other than diagnostic codes for plausible antidepressant treatment indications in the prediction models.[38] The IDI can be interpreted as the difference in discrimination slopes between the final and baseline models or alternatively as the change in average sensitivity (ie, the sensitivity averaged over all possible cut-off values between 0 and 1) minus the change in average "one minus specificity" when comparing the final model to the baseline model.[38]

Finally, we assessed model calibration (the accuracy of the predicted probabilities compared to the observed outcomes) by calculating the ratio of observed to expected number of prescriptions for indications other than depression within each of five strata based on the estimated probability of the outcome: 0–0.2, >0.2–0.4, >0.4–0.6, >0.6–0.8, and >0.8–1.0. The expected number of prescriptions for indications other than depression in each stratum was calculated by summing the estimated probabilities across all prescriptions in the corresponding stratum.

## Secondary outcome

As with the primary outcome, we assessed overall model performance for the secondary outcome using the scaled Brier score. However, because there were five treatment indication categories, we calculated the scaled Brier score using the following formula:

$$\text{scaled Brier score} = 1 - \left( \frac{1}{N} \sum_{j=1}^{R} \sum_{i=1}^{N} \left( \hat{Y}_{ij} - Y_{ij} \right)^2 \Big/ \frac{1}{N} \sum_{j=1}^{R} \sum_{i=1}^{N} \left( \bar{Y}_{ij} - Y_{ij} \right)^2 \right)$$

where $N$ represents the total number of antidepressant prescriptions, $R$ represents the number of treatment indication categories, $\hat{Y}_{ij}$ represents the estimated probability that prescription $i$ was written for indication $j$, $Y_{ij}$ represents the observed outcome for prescription $i$ corresponding to indication $j$ (ie, 1 if prescription $i$ was written for indication $j$, 0 otherwise), and $\bar{Y}_{ij}$ represent the overall (marginal) observed probability of indication $j$ in the study sample.[35] We also assessed the overall performance of the model separately for each treatment indication category using a "one-versus-rest" approach, where the scaled Brier score was calculated as per the primary outcome.

We calculated 95% confidence intervals (CIs) around all performance measures using a two-stage cluster bootstrap[39] to account for multi-level clustering of prescriptions within patients, who in turn were nested within physicians. The reported 95% CIs correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1,000 bootstrap re-samples of the test set.

## Measures of association

We used the final multivariable prediction model to estimate the adjusted odds ratio ($OR_{adj}$) for the independent association between each selected covariate and the odds of a treatment indication besides depression, fitted to the entire dataset (Figure 1, Step 5). We combined the training and test sets when estimating the adjusted odds ratios to maximize the precision of our estimates. As before, we used a two-stage cluster bootstrap[39] to calculate the 95% CIs around the $OR_{adj}$ estimates.

All statistical analyses were performed using SAS (SAS Institute Inc., Cary, NC, USA) software, version 9.4.

## Results

The study analysis included a total of 73,576 antidepressant prescriptions written by 141 physicians for 16,262 patients (Figure 1). Physicians prescribed antidepressants to a median of 70 (interquartile range [IQR] 12–171) patients,

and patients received a median of 3 (1–6) antidepressant prescriptions over the study period. Among all antidepressant prescriptions, 32,405 (44.0%) were for indications other than depression – 16,374 (22.3%) for anxiety/stress disorders, 7,295 (9.9%) for sleeping disorders, 4,137 (5.6%) for pain, and 4,599 (6.3%) for miscellaneous indications. The remaining 41,171 (56.0%) prescriptions were for depression.

## Prescription, patient, and physician characteristics

Overall, the most commonly prescribed antidepressants were venlafaxine (20.9%), citalopram (18.5%), and trazodone (9.4%). For only a small proportion (2.9%) of prescriptions, physicians prescribed the antidepressant on a "take-as-needed" basis.

More than two-thirds (68.9%) of all antidepressant prescriptions were written for female patients. Patients were a median of 55 (IQR 45–65) years old and 55.6% were registered in the public drug insurance plan. The percentage of patients without a diagnostic code for any of the 13 plausible antidepressant treatment indications within ±3 days of the index prescription date was 57.4%, which decreased to 22.5% when the length of the lookback window was increased to 365 days. Nearly one-third (31.7%) of patients had at least one chronic condition in the Charlson comorbidity index, 9.2% had been hospitalized in the past year, and 32.5% had visited the ER in the past year.

Nearly all antidepressant prescriptions were written by physicians who had received their medical training in Canada or the USA (90.0%) and had been practicing for at least 15 years at the start of the study period (94.0%). Physicians saw a median of 18 (IQR 13–23) patients per working day. Physician scores on the Evidence-Nonconformity-Practicality survey[33] suggested that they favored scientific evidence over clinical experience as the best source of knowledge (median evidence score of 20, IQR 19–22), were comfortable diverging from clinical norms and common practices (median nonconformity score of 17, IQR 15–19), and were not overly concerned with pragmatic concerns of practice (median practicality score of 15, IQR 14–17).

The distribution of most of these characteristics did not differ notably between prescriptions in the training and test sets (Table S6). The only exceptions were the proportion of antidepressants prescribed on a "take-as-needed" basis (3.8% in the training set compared to 0.8% in the test set) and the proportion of antidepressant prescriptions written by physicians trained in Canada or the USA (93.9% in the training set compared to 80.5% in the test set).

## Best FP1 function for continuous variables

Among all candidate predictors considered, 13 were continuous variables. For seven of these covariates, there was a better nonlinear FP1 function with a higher cross-validated scaled Brier score than the conventional linear $X^1$ representation. The best-fitting FP1 function was $X^{-2}$ for four of these variables (number of other drugs concurrently prescribed with the index drug, patient age, number of outpatient physicians seen in the past year, and the physician nonconformity score), $X^{-0.5}$ for two variables (number of outpatient visits in the past year and physician workload), and $X^3$ for the physician evidence score. The linear function $X^1$ was used for the remaining six continuous variables (Table 1).

## Derivation of the final model

Starting with the baseline model, which only contained the 26 variables based on diagnostic codes for plausible antidepressant treatment indications, the cross-validated estimate of the scaled Brier score was 0.0916 (Table 2). Thus, compared to a non-informative model where all prescriptions were assigned a probability of 44% (the overall probably that an antidepressant prescription was written for an indication other than depression), the baseline model reduced the mean square error by only 9.16%.

Among the remaining 347 candidate predictors, the forward stepwise selection procedure added 14 of these variables to the baseline model (Table 2). The name of the molecule prescribed was added first and was by far the best predictor of whether an antidepressant was prescribed for an indication other than depression. Adding this variable resulted in a cross-validated scaled Brier score of 0.3193 – representing an increase of 0.2277 over the baseline model. The next three variables added were the area-level marker of patient education, physician workload, and the prescribed dose, which collectively increased the cross-validated estimate of the scaled Brier score by another 0.0117. The last 10 variables added – each with only a very minor contribution to the cross-validated scaled Brier score – were the number of outpatient visits in the past year, whether the drug was prescribed on a "take-as-needed" basis, binary variables for whether the patient had been prescribed each of three drugs in the past year (trazodone, quetiapine, and furosemide), binary variables for whether the patient had a diagnostic code in the past year for each of three conditions (diabetes without chronic complication, dementia, and unspecified nonpsychotic mental disorder following organic brain damage), patient age, and a binary variable for whether the patient had a billing code for any diagnostic procedure (eg, endoscopies, electrocardiograms, biopsies) in the past year. Collectively adding these 14 variables to the baseline model increased the scaled Brier score by 0.2499, creating the final "main-terms" model with a scaled Brier score of 0.3415.

Finally, we tested whether the main-terms model performed better when a first-order interaction term was added

**Table 2** Derivation of the final prediction model for antidepressant prescriptions for indications other than depression

| Order added | Variables included in the model | Scaled Brier score$_{cv}$[a] | Δ[b] |
|---|---|---|---|
| - | 26 binary variables for the presence of diagnostic codes for plausible antidepressant treatment indications (the "baseline model") | 0.0916 | +0.0916 |
| 1 | Molecule name | 0.3193 | +0.2277 |
| 2 | Less than university education | 0.3233 | +0.0040 |
| 3 | Physician workload | 0.3274 | +0.0041 |
| 4 | Prescribed dose | 0.3310 | +0.0036 |
| 5 | Number of outpatient visits in the past year | 0.3327 | +0.0017 |
| 6 | Drug prescribed on a "take-as-needed" basis | 0.3342 | +0.0015 |
| 7 | Trazodone prescribed in the past year | 0.3357 | +0.0015 |
| 8 | Diagnostic code for diabetes without chronic complication in the past year | 0.3369 | +0.0011 |
| 9 | Diagnostic code for unspecified nonpsychotic mental disorder following organic brain damage (310.9) in the past year | 0.3380 | +0.0011 |
| 10 | Age | 0.3389 | +0.0009 |
| 11 | Any diagnostic procedure in the past year | 0.3397 | +0.0009 |
| 12 | Quetiapine prescribed in the past year | 0.3404 | +0.0007 |
| 13 | Furosemide prescribed in the past year | 0.3410 | +0.0006 |
| 14 | Diagnostic code for dementia in the past year ("main-terms" model) | 0.3415 | +0.0005 |
| 15 | Molecule name × prescribed dose (the "final model") | 0.3452 | +0.0037 |

**Notes:** [a]Cross-validated estimate of the scaled Brier score for predicting the primary outcome. Estimates were obtained using a 3-fold cross-validation procedure with the prescriptions in the training set. Higher scores indicate better overall model performance. [b]Change in the scaled Brier score when the corresponding variable was added to the previous model. The performance of the baseline model was compared to the performance of a non-informative model with no covariates, which by definition had a scaled Brier score of 0.

between the name of the molecule and the prescribed dose or when patient age was crossed with any variable for diagnostic codes or drugs prescribed in the past year. Only the interaction term between the name of the molecule and the prescribed dose further increased the cross-validated scaled Brier score, yielding the final prediction model with a cross-validated scaled Brier score of 0.3452 (Table 2). Table S7 shows the regression coefficients of all covariates in the final prediction model that were fit using the entire training set. These regression coefficients were used to calculate the estimated probability of the primary outcome for prescriptions in the test set.

## Performance of the final versus baseline models

### Primary outcome

In the test set, the final model had a scaled Brier score of 0.307 (95% CI 0.245 to 0.360) (Table 3). The final model had good discrimination, with a $c$ statistic of 0.815 (95% CI 0.787 to 0.847) and a discrimination slope of 0.325 (95% CI 0.286 to 0.366). The final model performed substantially better than the baseline model, which had a scaled Brier score of only 0.076 (95% CI –0.007 to 0.131), a $c$ statistic of only 0.651 (95% CI 0.590 to 0.711), and a discrimination slope of only 0.086 (95% CI 0.052 to 0.120). The IDI of 0.239 (95% CI 0.204 to 0.270) (Table 3) resulted from an increase in average sensitivity of 0.113 (95% CI 0.085 to 0.143) minus a decrease in average "one-minus-specificity" of –0.126 (95% CI –0.150 to –0.099), further demonstrating that the 14 variables added to the baseline model significantly improved the ability to discriminate between prescriptions written for depression versus other indications.

The overall calibration of the final model was very good, with a ratio of 0.986 (95% CI 0.842 to 1.136) for the overall number of observed to expected prescriptions for indications other than depression (Table 4). The final prediction model underestimated the number of prescriptions in the stratum with a very low (≤0.2)

probability of the outcome and slightly overestimated the number of prescriptions in the stratum with a higher probability of the outcome (0.2 to 0.6). When the estimated probability exceeded 0.6, however, the predictions from the final model were very accurate. In comparison, the overall calibration of the baseline model was worse, with a ratio of 0.935 (95% CI 0.773 to 1.125) for the overall number of observed to expected prescriptions for indications other than depression. Furthermore, the baseline model did not afford a clear identification of prescriptions with a high probability of indications other than depression, as over half of prescriptions had an estimated probability between 0.4 and 0.6 and only 0.3% (n=75) of prescriptions had an estimated probability >0.8 (Table 4).

### Secondary outcome

For the multinomial logistic regression model that produced estimated probabilities for each of the five treatment indication classes, the scaled Brier score across the five classes was 0.320 (95% CI 0.249 to 0.385) (Table 5). When the estimated probabilities for each treatment indication class were evaluated separately, the final multinomial model performed best for sleeping disorders (scaled Brier score of 0.628, 95% CI 0.518 to 0.736) and worst for miscellaneous indications (0.128, 95% CI 0.044 to 0.202). In comparison, the performance of the baseline model was again much worse. The scaled Brier score across the five indication categories was notably lower at only 0.067 (95% CI 0.002 to 0.108) and was also lower for each treatment indication class individually, especially sleeping disorders (scaled Brier score of only 0.029, 95% CI 0.004 to 0.043).

## Association between predictors in the final model and the primary outcome

Compared to venlafaxine (a serotonin-norepinephrine reuptake inhibitor [SNRI]), the molecules most likely to

**Table 3** Performance of the final and baseline models for predicting antidepressant prescriptions for indications other than depression

| | Performance in the test set[a] (95% CI) | | | |
| --- | --- | --- | --- | --- |
| | **Scaled Brier score[b]** | **c statistic** | **Discrimination slope[c]** | **IDI[d]** |
| Final model | 0.307 | 0.815 | 0.325 | 0.239 |
| (Diagnostic codes + other health-related information) | (0.245 to 0.360) | (0.787 to 0.847) | (0.286 to 0.366) | (0.204 to 0.270) |
| Baseline model | 0.076 | 0.651 | 0.086 | |
| (Diagnostic codes only) | (–0.007 to 0.131) | (0.590 to 0.711) | (0.052 to 0.120) | |

**Notes:** [a]Based on the regression coefficients for the final and baseline models that were fit using the training set and applied to the test set. [b]For a binary outcome, the scaled Brier score is analogous to the Pearson's $R^2$ statistic for continuous outcomes.[36] Higher scores indicate better performance. [c]Calculated as the absolute difference in the average probability of the outcome among observations with and without the outcome.[36] [d]Quantifies the incremental value of adding new markers to an existing model to predict a binary outcome. The IDI is equal to the difference in discrimination slopes between the final and baseline models, or alternatively, the difference between the change in average (ie, over all possible cut-off values between 0 and 1) sensitivity and the change in average "one minus specificity" when comparing the final model to the baseline model.[38]
**Abbreviation:** IDI, integrated discrimination improvement.

be prescribed for indications other than depression were amitriptyline (a tricyclic antidepressant [TCA]) (adjusted odds ratio [$OR_{adj}$] 20.98, 95% CI 12.27 to 48.91) and trazodone ($OR_{adj}$ 18.55, 95% CI 8.7 to 45.88) (Table 6). Other drugs more likely to be prescribed for indications besides depression than venlafaxine included four TCAs (nortriptyline, doxepin, imipramine, and desipramine), duloxetine (an SNRI), and paroxetine (a selective serotonin reuptake inhibitor [SSRI]). Bupropion was least likely to be prescribed for indications other than depression, with an $OR_{adj}$ of 0.18 (95% CI 0.06 to 0.44) compared to venlafaxine.

For certain molecules, indications besides depression were less likely if the prescribed dose increased. For each 10 mg/day increase in the prescribed dose, the odds of indications other than depression decreased substantially for mirtazapine ($OR_{adj}$ 0.68, 95% CI 0.48 to 0.89), nortriptyline ($OR_{adj}$ 0.68, 95% CI 0.30 to 0.92) and paroxetine ($OR_{adj}$ 0.78 (95% CI 0.66 to 0.91), and decreased moderately for citalopram ($OR_{adj}$ 0.86, 95% CI 0.76 to 0.96), amitriptyline ($OR_{adj}$

0.92, 95% CI 0.78 to 0.99) and venlafaxine ($OR_{adj}$ 0.96, 95% CI 0.94 to 0.98) (Table 6). On the other hand, antidepressants that were prescribed on a "take-as-needed" basis were notably more likely to be prescribed for indications other than depression ($OR_{adj}$ 2.85, 95% CI 1.47 to 6.09).

Antidepressants were also more likely to be prescribed for indications other than depression if the patient had undergone a diagnostic procedure in the past year ($OR_{adj}$ 1.19, 95% CI 1.04 to 1.33) or lived in an area where a higher percentage of the population did not have university education ($OR_{adj}$ per 1% increase 1.07, 95% CI 1.03 to 1.10). Patients with a diagnostic code for anxiety/stress disorders, fibromyalgia, obsessive-compulsive disorder, or attention deficit/hyperactivity disorder in the past year were also more likely to be prescribed antidepressants for indications other than depression, with these associations being strongest when the diagnostic code was recorded around the index prescription date. Conversely, patients were significantly less likely to be prescribed antidepressants for indications other than depres-

**Table 4** Calibration of the final and baseline models for predicting antidepressant prescriptions for indications other than depression

| Probability of treatment indication other than depression[a] | Performance in the test set | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Final model (diagnostic codes + other health-related information) | | | | Baseline model (diagnostic codes only) | | | |
| | N | O | E[b] | O:E (95% CI) | N | O | E[b] | O:E (95% CI) |
| 0–0.2 | 5,531 | 756 | 571.15 | 1.324 (0.911 to 1.701) | 2,398 | 424 | 385.67 | 1.099 (0.825 to 1.514) |
| >0.2–0.4 | 5,646 | 1,551 | 1,703.99 | 0.910 (0.659 to 1.215) | 3,497 | 936 | 1,070.39 | 0.874 (0.733 to 1.054) |
| >0.4–0.6 | 4,427 | 2,088 | 2,171.00 | 0.962 (0.730 to 1.243) | 11,538 | 5,331 | 5,765.59 | 0.925 (0.736 to 1.158) |
| >0.6–0.8 | 2,254 | 1,521 | 1,544.70 | 0.985 (0.833 to 1.152) | 4,049 | 2,515 | 2,624.67 | 0.958 (0.792 to 1.134) |
| >0.8–1.0 | 3,699 | 3,357 | 3,417.19 | 0.982 (0.931 to 1.030) | 75 | 67 | 66.11 | 1.013 (0.768 to 1.133) |
| Overall | 21,557 | 9,273 | 9,408.03 | 0.986 (0.842 to 1.136) | 21,557 | 9,273 | 9,912.43 | 0.935 (0.773 to 1.125) |

**Notes:** [a]The probability of the outcome was calculated for prescriptions in the test set based on the regression coefficients obtained using the training set. [b]The expected number of prescriptions for a treatment indication besides depression was calculated by summing the probabilities across all prescriptions in the stratum.
**Abbreviations:** N, number of antidepressant prescriptions; O, observed number of antidepressant prescriptions for a treatment indication besides depression; E, expected number of antidepressant prescriptions for a treatment indication besides depression; O:E, ratio of observed to expected prescriptions.

**Table 5** Overall and per-class performance of the final and baseline models for predicting antidepressant treatment indications expressed as a five-class outcome

| Treatment indication class | Performance in the test set | |
|---|---|---|
| | Scaled Brier score[a] (95% CI) | |
| | Final model (diagnostic codes + other health-related information) | Baseline model (diagnostic codes only) |
| Depression | 0.312 (0.255 to 0.371) | 0.075 (–0.018 to 0.131) |
| Anxiety/stress disorders | 0.223 (0.122 to 0.297) | 0.084 (–0.004 to 0.146) |
| Sleeping disorders | 0.628 (0.518 to 0.736) | 0.029 (0.004 to 0.043) |
| Pain | 0.356 (0.041 to 0.556) | 0.042 (–0.024 to 0.079) |
| Miscellaneous | 0.128 (0.044 to 0.202) | 0.057 (0.011 to 0.100) |
| All indications | 0.320 (0.249 to 0.385) | 0.067 (0.002 to 0.108) |

**Note:** [a]Based on the regression coefficients for the final and baseline models that were fit using the training set and applied to the test set. The per-class estimates were calculated using a one-versus-rest approach.

**Table 6** Adjusted odds ratios for the independent association between variables in the final prediction model and antidepressant prescriptions for treatment indications other than depression

| | Nᵃ (%) or median (IQR) | Antidepressant prescriptions for treatment indications other than depression | |
|---|---|---|---|
| | | Adjusted ORᵇ | 95% CI |
| **Prescription-related factors** | | | |
| Molecule name | | | |
| Venlafaxineᵈ | 15,398 (20.9) | 1.00 | [Reference] |
| Amitriptylineᵉ | 6,196 (8.4) | 20.98 | 12.27 to 48.91 |
| Trazodone | 6,891 (9.4) | 18.55 | 8.7 to 45.88 |
| Nortriptylineᵉ | 434 (0.6) | 16.32 | 5.43 to 190.13 |
| Doxepinᵉ | 461 (0.6) | 10.49 | 2.60 to 109.35 |
| Imipramineᵉ | 200 (0.3) | 8.84 | 1.39 to 301.56 |
| Desipramineᵉ | 138 (0.2) | 3.98 | 1.31 to 73.55 |
| Duloxetineᵈ | 1,596 (2.2) | 2.40 | 1.10 to 6.10 |
| Paroxetineᶜ | 6,751 (9.2) | 2.05 | 1.11 to 3.64 |
| Clomipramineᵉ | 165 (0.2) | 1.54 | 0.26 to 12.36 |
| Citalopramᶜ | 13,623 (18.5) | 1.07 | 0.67 to 1.69 |
| Escitalopramᶜ | 4,470 (6.1) | 0.82 | 0.53 to 1.51 |
| Sertralineᶜ | 4,457 (6.1) | 0.74 | 0.45 to 1.26 |
| Fluvoxamineᶜ | 669 (0.9) | 0.72 | 0.20 to 1.50 |
| Trimipramineᵉ | 436 (0.6) | 0.69 | 0.20 to 3.01 |
| Fluoxetineᶜ | 1,451 (2.0) | 0.65 | 0.27 to 1.50 |
| Mirtazapine | 4,132 (5.6) | 0.45 | 0.18 to 1.02 |
| Bupropion | 5,631 (7.7) | 0.18 | 0.06 to 0.44 |
| Desvenlafaxineᵈ | 477 (0.7) | 0.18 | 0.02 to 310,670.11 |
| Prescribed dose (mg/day), per 10 mg increase by molecule | | | |
| Mirtazapine | 30 (15–30) | 0.68 | 0.48 to 0.89 |
| Nortriptylineᵉ | 25 (10–50) | 0.68 | 0.30 to 0.92 |
| Paroxetineᶜ | 20 (15–30) | 0.78 | 0.66 to 0.91 |
| Desvenlafaxineᵈ | 50 (50–100) | 0.83 | 0.05 to 1.13 |
| Doxepinᵉ | 40 (25–75) | 0.85 | 0.62 to 1.07 |
| Citalopramᶜ | 20 (20–30) | 0.86 | 0.76 to 0.96 |
| Imipramineᵉ | 50 (25–75) | 0.86 | 0.62 to 1.11 |
| Fluoxetineᶜ | 20 (20–40) | 0.87 | 0.65 to 1.11 |
| Desipramineᵉ | 50 (25–100) | 0.90 | 0.52 to 1.07 |
| Amitriptylineᵉ | 20 (10–30) | 0.92 | 0.78 to 0.99 |
| Escitalopramᶜ | 10 (10–20) | 0.95 | 0.64 to 1.14 |
| Duloxetineᵈ | 60 (30–60) | 0.96 | 0.80 to 1.11 |
| Venlafaxineᵈ | 75 (75–150) | 0.96 | 0.94 to 0.98 |
| Trimipramineᵉ | 50 (25–75) | 0.96 | 0.64 to 1.10 |
| Sertralineᶜ | 50 (50–100) | 0.99 | 0.94 to 1.02 |
| Trazodone | 50 (50–100) | 0.99 | 0.95 to 1.05 |
| Fluvoxamineᶜ | 100 (50–143) | 1.00 | 0.90 to 1.09 |
| Clomipramineᵉ | 75 (30–100) | 1.01 | 0.79 to 1.24 |
| Bupropion | 150 (150–300) | 1.02 | 0.99 to 1.06 |
| Drug prescribed on a "take-as-needed" basis | 2,117 (2.9) | 2.85 | 1.47 to 6.09 |
| **Patient-related factors** | | | |
| Any diagnostic procedure in the past year | 24,542 (33.4) | 1.19 | 1.04 to 1.33 |
| Less than university education (%), per 1% increase | 19.2 (16.8–20.6) | 1.07 | 1.03 to 1.10 |
| Diagnostic codes in the past year | | | |
| Unspecified nonpsychotic mental disorder following organic brain damage (310.9) | 980 (1.3) | 0.48 | 0.26 to 0.85 |
| Dementia | 1,085 (1.5) | 0.74 | 0.49 to 1.09 |
| Diabetes without chronic complication | 8,197 (11.1) | 0.82 | 0.67 to 1.00 |

*(Continued)*

**Table 6** (*Continued*)

| Plausible antidepressant treatment indications[f] | N[a] (%) or median (IQR) | | Antidepressant prescriptions for treatment indications besides depression | | | |
| | | | Adjusted OR[b] | | 95% CI | |
| | ± 3 days | −4 to −365 days | ± 3 days | | | −4 to −365 days |
|---|---|---|---|---|---|---|
| Depression | 13,600 (18.5) | 22,028 (29.9) | 0.40 | 0.31 to 0.49 | 0.46 | 0.36 to 0.56 |
| Anxiety/stress disorders | 11,106 (15.1) | 22,192 (30.2) | 2.09 | 1.61 to 2.71 | 1.52 | 1.27 to 1.89 |
| Sleeping disorders | 681 (0.9) | 3,314 (4.5) | 1.55 | 0.97 to 2.40 | 0.99 | 0.79 to 1.26 |
| Pain | 3,881 (5.3) | 25,392 (34.5) | 1.22 | 1.02 to 1.48 | 1.01 | 0.92 to 1.12 |
| Migraine | 684 (0.9) | 3,891 (5.3) | 1.33 | 0.83 to 2.04 | 0.93 | 0.77 to 1.12 |
| Fibromyalgia | 775 (1.1) | 2,640 (3.6) | 2.21 | 1.43 to 3.47 | 1.44 | 1.04 to 2.08 |
| Obsessive-compulsive disorder | 169 (0.2) | 349 (0.5) | 14.53 | 4.68 to 134.36 | 3.59 | 1.84 to 7.10 |
| Vasomotor symptoms of menopause | 562 (0.8) | 2,787 (3.8) | 1.34 | 0.82 to 2.17 | 1.16 | 0.90 to 1.52 |
| Nicotine dependence | 106 (0.1) | 458 (0.6) | 2.26 | 0.80 to 5.20 | 1.05 | 0.58 to 1.72 |
| Attention deficit/hyperactivity disorder | 114 (0.2) | 387 (0.5) | 2.51 | 1.03 to 6.92 | 1.37 | 0.65 to 2.57 |
| Sexual dysfunction | 10 (0.0) | 95 (0.1) | 1.58 | 0.0 to 91,807.54 | 1.41 | 0.49 to 3.75 |
| Pre-menstrual dysphoric disorder | 26 (0.0) | 82 (0.1) | 1.77 | 0.22 to 80,903.71 | 0.90 | 0.29 to 3.54 |
| Eating disorders | 31 (0.0) | 145 (0.2) | 2.28 | 0.39 to 29.76 | 2.02 | 0.70 to 4.86 |
| Drugs prescribed in the past year | | | | | | |
| Furosemide | 1,896 (2.6) | | 0.62 | | 0.37 to 0.98 | |
| Trazodone | 7,175 (9.8) | | 0.71 | | 0.54 to 0.92 | |
| Quetiapine | 4,100 (5.6) | | 0.77 | | 0.58 to 1.04 | |

**Notes:** [a]Total N = 73,576. [b]Adjusted ORs were obtained using the regression coefficients from a multivariable logistic regression model that were fit using all prescriptions (ie, training and test sets combined). Adjusted ORs for the three continuous covariates that were expressed using nonlinear FP1 functions are not shown in this table. [c]Selective serotonin reuptake inhibitors. [d]Serotonin-norepinephrine reuptake inhibitors. [e]Tricyclic antidepressants. [f]The presence of diagnostic codes for the 13 plausible antidepressant treatment indication categories was examined in two separate observation windows: ±3 days and −4 to −365 days.
**Abbreviations:** N, number of antidepressant prescriptions; IQR, interquartile range.

sion if they had a diagnostic code for depression, unspecified nonpsychotic mental disorder following organic brain damage (ICD-9 code of 310.9), or diabetes without chronic complication, or if they had been prescribed furosemide or trazodone in the past year. Patients with diagnostic codes for dementia or prescriptions for quetiapine in the past year were also less likely to be prescribed antidepressants for indications other than depression, but these 95% CIs did not exclude 1 (Table 6).

Figure 2 shows the adjusted ORs for the primary outcome that were estimated in the final model for the three continuous covariates that were expressed using nonlinear FP1 functions: patient age (panel A), the number of outpatient visits in the past year (panel B), and physician workload (panel C). For all these covariates, the odds of the outcome decreased with increasing values of the variable over their low to middle ranges but plateaued for values above a threshold.

## Discussion

In this study, we derived and validated a model that could accurately predict when primary care physicians prescribed antidepressants for indications other than depression. The most important predictors in the final model included the name of the molecule and the dose at which it was prescribed, the presence of diagnostic codes for certain conditions in administrative data over the past year, the patient's education level, and the physician's workload. The final prediction model had good discrimination, good calibration, and performed substantially better than a model containing only covariates based on diagnostic codes for plausible antidepressant treatment indications.

## Comparison with previous studies

Few studies have attempted to predict antidepressant treatment indications or identify factors associated with them. Gardarsdottir et al[25] developed an algorithm to identify antidepressant users with a diagnostic code for depression in a Dutch medical database. The authors similarly found that antidepressant users with a diagnostic code for depression were more likely to be prescribed SSRIs rather than TCAs and to be prescribed higher doses of the drug. Milea et al[15] identified factors associated with antidepressant users without a diagnostic code for either approved or common off-label indications in a US claims database. Although the authors dissimilarly found that antidepressant users without a diagnosis of interest were more likely to be older or female, it is
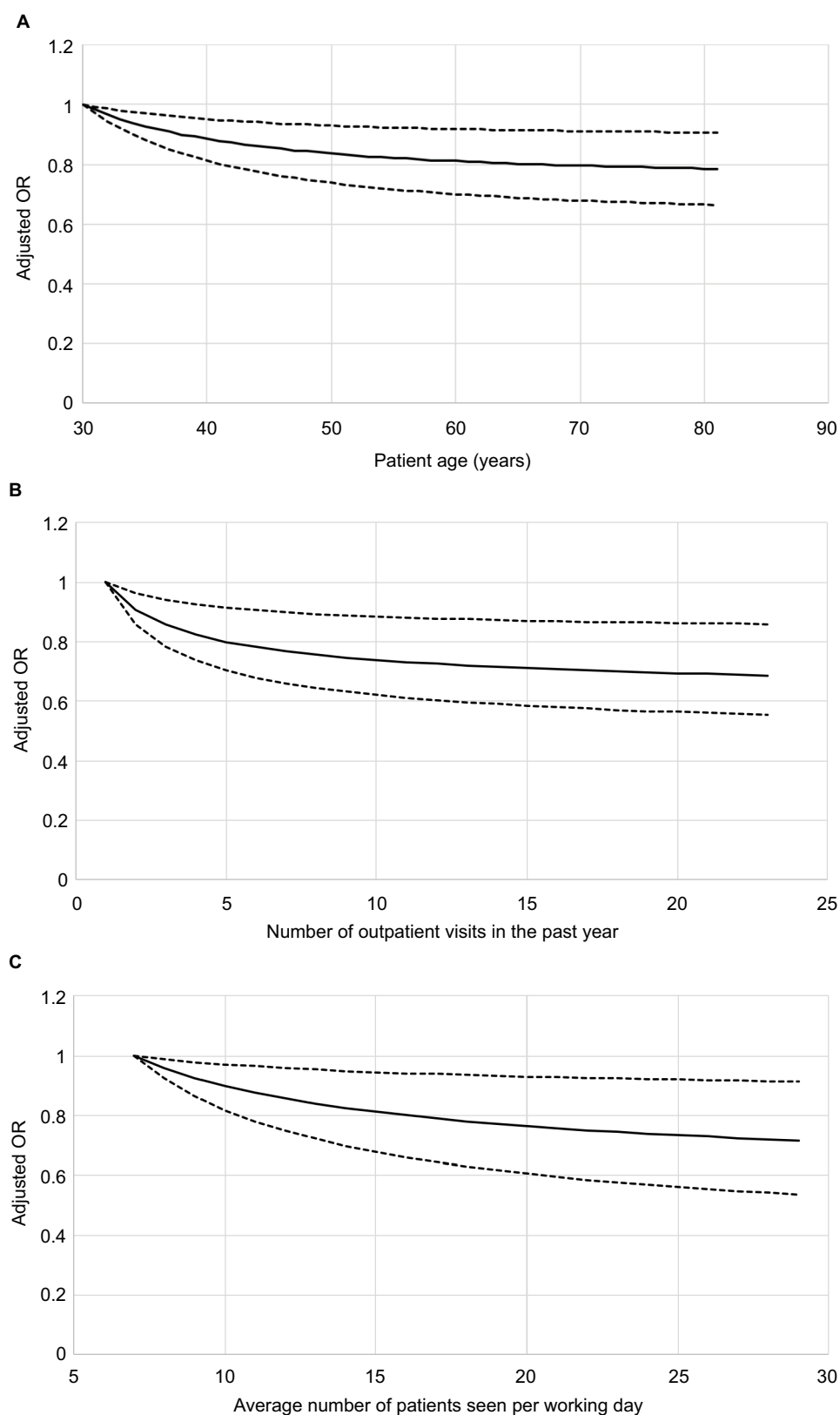
**A**



**B**



**C**



**Figure 2** Independent association between antidepressant prescriptions for indications other than depression and the three continuous covariates in the final model that were expressed using non-linear FP1 functions. Patient age (**A**) was expressed using the function $X^{-2}$ while the number of outpatient visits in the past year (**B**) and physician workload (**C**) were expressed using the function $X^{-0.5}$. The adjusted ORs account for all other covariates in the final model and were calculated based on coefficients fit using all prescriptions. For each continuous covariate, adjusted ORs were calculated from the 5th to 95th percentile of the distribution of observed values using the value at the 5th percentile as the reference level. The black lines represent the point estimates of the adjusted ORs, while the dotted lines represent the 95% CIs around the point estimates.

hard to compare our findings because the authors combined depression with other off-label indications for antidepressants and did not include prescriptions for TCAs or trazodone in the analysis. Finally, Sihvo et al[40] identified factors associated with non-psychiatric antidepressant use (neither depression nor anxiety) in an adult Finnish population where patients' psychiatric history was assessed via self-report in a structured interview. The authors found that patients with lower levels of education or who had not used health services in the past year were more likely to take antidepressants for non-psychiatric reasons. Although neither of these associations was statistically significant in their study, these findings are concordant with our observation that patients with less than university education or fewer outpatient visits in the past year were more likely to receive antidepressant prescriptions for indications other than depression.

In all these studies, treatment indications were measured using suboptimal methods. The use of diagnostic codes alone identifies the true indication for only a small proportion of antidepressant prescriptions[19] and the use of self-reported psychiatric history may lead to under-reporting of depression and other mental conditions.[40] In this study, we measured treatment indications using validated, physician-documented indications that were systematically recorded for every prescription. We also considered a more extensive range of predictors for antidepressant treatment indications than any other previous studies.

## Explanations for study findings

Our finding that patients with lower education were more likely to be prescribed antidepressants for indications besides depression may be because patients with lower education are more likely to suffer from insomnia[41] and chronic pain[42] for which antidepressants are often prescribed. Our finding that antidepressant prescriptions were more likely to be written for depression if the patient had more previous outpatient visits may be explained by the observation that patients with depression visit their primary care provider more frequently than patients without depression.[43] We also found that patients with past prescriptions for furosemide were more likely to be prescribed antidepressants for depression, which may have been observed because depression is associated with heart failure and adverse renal disease outcomes.[44] Our finding that patients with previous prescriptions for trazodone were more likely to be prescribed antidepressants for depression may be due to the fact that sleeping disorders are a key symptom of depression[45] and the physicians in our study mainly prescribed trazodone for sleeping disorders (81%

of all trazodone prescriptions). Finally, the somatic and symptomatic nature of many nonpsychiatric conditions for antidepressants (eg, pain, fibromyalgia, Crohn's disease) may require patients to undergo various diagnostic tests before reaching a proper diagnosis.[46,47] Thus, the work-up required to investigate patients' medical complaints may explain why antidepressants were more likely to be prescribed for indications besides depression if the patient had undergone a diagnostic procedure in the past year.

## Implications of findings

Administrative data and electronic health records rarely contain information linking prescriptions directly to their corresponding diagnoses,[15,48] thus creating major barriers for using these data to study antidepressant use by indication. The fact that we derived a model that could accurately predict when antidepressants were being prescribed for indications other than depression is a promising finding for researchers. Specifically, it suggests that in the absence of documented treatment indications, researchers may still be able to use health services data to stratify antidepressant users by indication to achieve better precision when studying outcomes from antidepressant use and using antidepressant prescriptions as a proxy for depression treatment. Thus, our prediction model represents a valuable tool for enabling more valid, accurate database research on antidepressant use for both depression and other (eg, off-label) indications. Furthermore, the poor performance of our baseline model emphasizes that algorithms based on diagnostic codes alone should not be used to infer antidepressant treatment indications. Such algorithms likely have poor accuracy and will misclassify a significant number of antidepressant users, thus compromising the validity of the analysis.

For policy makers, the fact that the specific molecule prescribed was by far the strongest predictor of why antidepressants were prescribed suggests that health policies or interventions aimed at specific drugs may have large impacts on changing prescribing behaviors for antidepressants. For example, policies around trazodone would likely have a substantial impact on antidepressant use for sleeping disorders since this drug is used almost exclusively to treat insomnia.

## Strengths and limitations

The main strength of this study is that it included ten years of data where physicians systematically documented treatment indications for antidepressants at the point of prescribing. Another strength is that in deriving our prediction model, we applied sound practices to reduce model over-fitting

while still optimizing predictive performance. First, we used changes in the scaled Brier score rather than *p*-values as our criterion for selecting variables. Given that there were so many candidate predictors, the standard practice of relying on *p*-values[15,25,40] would have likely included many irrelevant variables or variables with little predictive value. The scaled Brier score, on the other hand, allowed us to assess each variable based on its predictive utility rather than statistical significance. Second, using 3-fold cross-validation during the variable selection process allowed us to obtain better estimates of the test error, thus reducing the risk of overfitting the final model to the training set. Finally, we tested the final model's performance on a held-out set of prescriptions that had not been used during the training process, which allowed us to better estimate the model's performance on new data. However, despite our best efforts to prevent overfitting, the final prediction model still performed better in the training set (cross-validated estimate of 0.345 for the scaled Brier score) than in the test set (scaled Brier score of 0.307).

Study limitations include the generalizability of our findings, as the MOXXI system is used by primary care physicians in one Canadian province for patients that are generally older with more health complexities.[24] Future research should be done to determine if our findings represent the prescribing habits of other specialists and physicians from other geographical locations. Another study limitation is that in creating predictors for the model, we did not consider previously dispensed drugs, because dispensing data was unavailable for patients with private drug insurance who contributed nearly half (44%) of the antidepressant prescriptions in this study. We also modelled treatment indications as mutually exclusive categories even though some antidepressants could have been prescribed for multiple indications. However, we hypothesize that this situation did not occur very frequently because only a small proportion (1.8%) of antidepressant prescriptions had multiple indications entered in the electronic prescribing system. This observation is similar to another UK study[49] that found only 5% of antidepressants users reported taking antidepressants for multiple indications. Furthermore, MOXXI physicians had good reason to enter all relevant indications into the MOXXI system because these conditions became part of the patient's problem list, which was then accessed by drug knowledge databases to identify potential drug-disease problems.[22]

Finally, two study considerations deserve mention. First, because we used prescribing data in this study, future research should determine whether the performance of our prediction model is similar when applied to drug dispensing data, since patients who fill their antidepressant prescriptions may be distinct from those who do not. Second, the relationships we observed in this study were associational and not necessarily causal, thus requiring further investigation to reveal the mechanisms behind them.

## Conclusion

In this study, we used health services data to derive a model that could accurately predict when primary care physicians prescribed antidepressants for indications other than depression. This model represents a valuable tool that could enable researchers to conduct more accurate database studies on antidepressant use by indication in the absence of documented treatment indications. This study also identified important factors associated with antidepressant prescriptions for indications besides depression that could help inform health policies and interventions aimed at changing prescribing behaviors for antidepressants.

## Acknowledgment

## Disclosure

## References

1. Shepherd G, Mohorn P, Yacoub K, May DW. Adverse drug reaction deaths reported in United States vital statistics, 1999–2006. *Ann Pharmacother*. 2012;46(2):169–175.
2. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*. 1998;279(15):1200–1205.
3. Bonn D. Adverse drug reactions remain a major cause of death. *Lancet*. 1998;351(9110):1183.
4. Wu C, Bell CM, Wodchis WP. Incidence and economic burden of adverse drug reactions among elderly patients in Ontario emergency departments: a retrospective study. *Drug Saf*. 2012;35(9):769–781.
5. World Health Organization. *The Importance of Pharmacovigilance: Safety Monitoring of Medicinal Products*. Geneva: World Health Organization; 2002.
6. Kumar A. Pharmacovigilance: importance, concepts, and processes. *Am J Health Syst Pharm*. 2017;74(8): 606–612.
7. Pratt L, Brody D, Gu Q. Antidepressant use in persons aged 12 and over: United States, 2005–2008. *NCHS Data Brief*, 2011;76:1–8. Available from: http://www.cdc.gov/nchs/data/databriefs/db76.htm. Accessed May 5, 2015.

8. Rotermann M, Sanmartin C, Hennessy D, Arthur M. Prescription medication use by Canadians aged 6 to 79. *Health Rep*. 2014;25(6):3–9. Available from: http://www.statcan.gc.ca/pub/82-003-x/2014006/article/14032-eng.htm. Accessed May 5, 2015.

9. Prescribing and Medicines Team. *Prescriptions Dispensed in the Community: England 2005–2015*. Health and Social Care Information Centre; 2016. Available from: http://digital.nhs.uk/catalogue/PUB20664/pres-disp-com-eng-2005-15-rep.pdf. Accessed September 14, 2016.

10. Wong J, Motulsky A, Eguale T, Buckeridge DL, Abrahamowicz M, Tamblyn R. Treatment indications for antidepressants prescribed in primary care in Quebec, Canada, 2006–2015. *JAMA*. 2016;315(20):2230–2232.

11. Wong J, Motulsky A, Abrahamowicz M, Eguale T, Buckeridge DL, Tamblyn R. Off-label indications for antidepressants in primary care: descriptive study of prescriptions from an indication based electronic prescribing system. *BMJ*. 2017;356:j603.

12. Eguale T, Buckeridge DL, Verma A, et al. Association of off-label drug use and adverse drug events in an adult population. *JAMA Intern Med*. 2016;176(1):55–63.

13. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects--advantages and disadvantages. *Nat Clin Pract Rheumatol*. 2007;3(12):725–732.

14. Larson MJ, Miller K, Fleming KJ. Treatment with antidepressant medications in private health plans. *Adm Policy Ment Health*. 2007;34(2):116–126.

15. Milea D, Verpillat P, Guelfucci F, Toumi M, Lamure M. Prescription patterns of antidepressants: findings from a US claims database. *Curr Med Res Opin*. 2010;26(6):1343–1353.

16. Simon GE, Stewart C, Beck A, et al. National prevalence of receipt of antidepressant prescriptions by persons without a psychiatric diagnosis. *Psychiatr Serv*. 2014;65(7):944–946.

17. Chen H, Reeves JH, Fincham JE, Kennedy WK, Dorfman JH, Martin BC. Off-label use of antidepressant, anticonvulsant, and antipsychotic medications among Georgia Medicaid enrollees in 2001. *J Clin Psychiatry*. 2006;67(6):972–982.

18. Wu C-S, Shau W-Y, Chan H-Y, Lee Y-C, Lai Y-J, Lai M-S. Utilization of antidepressants in Taiwan: a nationwide population-based survey from 2000 to 2009. *Pharmacoepidemiol Drug Saf*. 2012;21(9):980–988.

19. Wong J, Abrahamowicz M, Buckeridge DL, Tamblyn R. Assessing the accuracy of using diagnostic codes from administrative data to infer antidepressant treatment indications. *Pharmacoepidemiol Drug Saf*. In press.

20. Régie de l'assurance maladie du Québec. *Medical services*. Available from: http://www.ramq.gouv.qc.ca/en/citizens/health-insurance/health-care/Pages/medical-services.aspx. Accessed March 11, 2017.

21. Monfared AAT, Lelorier J. Accuracy and validity of using medical claims data to identify episodes of hospitalizations in patients with COPD. *Pharmacoepidemiol Drug Saf*. 2006;15(1):19–29.

22. Tamblyn R, Huang A, Kawasumi Y, et al. The development and evaluation of an integrated electronic prescribing and drug management system for primary care. *J Am Med Inform Assoc*. 2006;13(2):148–159.

23. Eguale T, Winslade N, Hanley JA, Buckeridge DL, Tamblyn R. Enhancing pharmacosurveillance with systematic collection of treatment indication in electronic prescribing: a validation study in Canada. *Drug Saf*. 2010;33(7):559–567.

24. Bartlett G, Tamblyn R, Kawasumi Y, Poissant L, Taylor L. Non-participation bias in health services research using data from an integrated electronic prescribing project: The role of informed consent. *Acta Bioethica*. 2005;11(2):145–159.

25. Gardarsdottir H, Egberts ACG, van Dijk L, Sturkenboom MCJM, Heerdink ER. An algorithm to identify antidepressant users with a diagnosis of depression from prescription data. *Pharmacoepidemiol Drug Saf*. 2009;18(1):7–15.

26. Stone KJ, Viera AJ, Parman CL. Off-label applications for SSRIs. *Am Fam Physician*. 2003;68(3):498–504.

27. McCall WV. Off-label use of prescription medications for insomnia: sedating antidepressants, antipsychotics, anxiolytics, and anticonvulsants (Chapter 34). In: *Insomnia: Diagnosis and Treatment*. London, UK: Informa Healthcare; 2010:397–409.

28. Canadian Institute for Health Information. Conversion Tables (for use with ICD-10-CA/CCI). https://secure.cihi.ca/estore/productSeries.htm?pc=PCC85. Accessed July 27, 2016.

29. Katon WJ. Epidemiology and treatment of depression in patients with chronic medical illness. *Dialogues Clin Neurosci*. 2011;13(1):7–23.

30. Druss BG, Rask K, Katon WJ. Major Depression, Depression Treatment, and Quality of Primary Medical Care. *Gen Hosp Psychiatry*. 2008;30(1):20–25.

31. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130–1139.

32. Papakostas GI. Managing partial response or nonresponse: switching, augmentation, and combination strategies for major depressive disorder. *J Clin Psychiatry*. 2009;70 Suppl 6:16–25.

33. Green LA, Gorenflo DW, Wyszewianski L, Michigan Consortium for Family Practice Research. Validating an instrument for selecting interventions to change physician practice patterns: a Michigan Consortium for Family Practice Research study. *J Fam Pract*. 2002;51(11):938–942.

34. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer-Verlag; 2009.

35. Brier GW. Verification of forecasts expressed in terms of probability. *Mon Weather Rev*. 1950;78(1):1–3.

36. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010;21(1):128–138.

37. Sauerbrei W, Meier-Hirmer C, Benner A, Royston P. Multivariable regression model building by using fractional polynomials: description of SAS, STATA and R programs. *Comput Stat Data Anal*. 2006;50(12):3464–3485.

38. Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157–172.

39. Xiao Y, Abrahamowicz M. Bootstrap-based methods for estimating standard errors in Cox's regression analyses of clustered event times. *Stat Med*. 2010;29(7–8):915–923.

40. Sihvo S, Isometsä E, Kiviruusu O, et al. Antidepressant utilisation patterns and determinants of short-term and non-psychiatric use in the Finnish general adult population. *J Affect Disord*. 2008;110(1–2):94–105.

41. Sutton DA, Moldofsky H, Badley EM. Insomnia and health problems in Canadians. *Sleep*. 2001;24(6):665–670.

42. Smith BH, Elliott AM, Chambers WA, Smith WC, Hannaford PC, Penny K. The impact of chronic pain in the community. *Fam Pract*. 2001;18(3):292–299.

43. Wong ST, Manca D, Barber D, et al. The diagnosis of depression and its treatment in Canadian primary care practices: an epidemiological study. *CMAJ Open*. 2014;2(4):E337–E342.

44. Kop WJ, Seliger SL, Fink JC, et al. Longitudinal association of depressive symptoms with rapid kidney function decline and adverse clinical renal disease outcomes. *Clin J Am Soc Nephrol*. 2011;6(4):834–844.

45. Nutt D, Wilson S, Paterson L. Sleep disorders as core symptoms of depression. *Dialogues Clin Neurosci*. 2008;10(3):329–336.

46. Berger A, Dukes E, Martin S, Edelsberg J, Oster G. Characteristics and healthcare costs of patients with fibromyalgia syndrome. *Int J Clin Pract*. 2007;61(9):1498–1508.

47. Wilkins T, Jarvis K, Patel J. Diagnosis and management of Crohn's disease. *Am Fam Physician*. 2011;84(12):1365–1375.

48. Li Y, Salmasian H, Harpaz R, Chase H, Friedman C. Determining the reasons for medication prescriptions in the EHR using knowledge and natural language processing. *AMIA Annu Symp Proc*. 2011;2011:768–776.

49. Aarts N, Noordam R, Hofman A, Tiemeier H, Stricker BH, Visser LE. Self-reported indications for antidepressant use in a population-based cohort of middle-aged and elderly. *Int J Clin Pharm*. 2016;38(5):1311–1317.