Using information from an indication-based electronic prescribing system to increase pharmacovigilance around antidepressant use for indications besides depression

Jenna CL Wong

Department of Epidemiology, Biostatistics, and Occupational Health

McGill University, Montreal, Canada

August 2017

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Epidemiology

Abstract

Background

The World Health Organization describes pharmacovigilance as “the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems.” An important component of pharmacovigilance is monitoring the use of drugs for medical reasons (indications) that are not approved (off-label) and lack supporting scientific evidence. Antidepressants are a prime example of drugs in need of heightened pharmacovigilance. These drugs are among the most commonly used medications in North America and are prescribed to treat not only depression, but also a host of other indications, many of which are off-label and lack supporting evidence. The ability to carry out pharmacovigilance activities on antidepressant use for different indications is severely hampered by the fact that treatment indications for medications are not routinely documented. Moreover, there are currently no validated mechanisms that exist to predict this missing information. Data from indication-based electronic prescribing systems could address these shortcomings and be used to conduct important pharmacovigilance activities on antidepressants. However, because such systems have not yet been widely implemented, the creation of validated algorithms that can accurately predict when antidepressants are being prescribed for indications besides depression could be valuable tools for enabling pharmacovigilance activities on antidepressants, as well as other medications that may be used for multiple treatment indications.

Objectives

The overall goal of this thesis was to increase the capacity for comprehensive pharmacovigilance around antidepressants by using data from a unique indication-based electronic prescribing system to address four specific research objectives:

1) Determine the prevalence of different treatment indications, including off-label indications, for antidepressants in primary care and to assess the level of scientific support for off-label indications;

2) Evaluate the accuracy of using diagnostic codes recorded in administrative health data to infer treatment indications, compared to gold-standard treatment indications for antidepressants;
3) Use standard regression techniques to predict and identify important determinants of antidepressant prescriptions for indications besides depression; and
4) Explore the use of more flexible machine-learning algorithms and an ensemble learning approach called "super learning" to predict when antidepressant prescriptions are written for indications besides depression.

Data sources

These four research objectives were addressed in five manuscripts. The first two manuscripts addressed the first research objective, while the last three manuscripts addressed the last three research objectives. For all these manuscripts, the source of data was the Medical Office of the XXIst Century (MOXXI) – an indication-based electronic prescribing and drug management system used by approximately 185 primary care physicians in Quebec since 2003. For all patients who were prescribed antidepressants in the MOXXI system, additional data on their use of health services in the past year were obtained by linking these patients to administrative health datasets from the provincial health insurance agency (the Régie de l'assurance maladie du Québec [RAMQ]) and the provincial hospitalization discharge summary databases (MED-ECHO).

Methods and Results

Manuscript 1: In this descriptive paper, I measured the prevalence of different treatment indications for antidepressant prescriptions that were written by primary care physicians in Quebec between January 1, 2006 and September 30, 2015. I classified treatment indications first by clinical condition, and then by label status (i.e. approved or not approved). I also measured changes over time in antidepressant prescribing for depression by using generalized linear risk difference models to estimate the linear effect of calendar time (in years) on the probability of antidepressant prescriptions being written for depression. Overall, I found that

only 55.2% of all antidepressant prescriptions written during the study period were for depression. Physicians also commonly prescribed antidepressants for anxiety disorders (18.5%), insomnia (10.2%), pain (6.1%), and panic disorders (4.1%). Between 2006 and 2015, the proportion of antidepressant prescriptions for depression decreased significantly for all major classes of antidepressants, especially serotonin-norepinephrine reuptake inhibitors (five-year risk difference of -9.73%; 95% CI, -11.86% to -7.61%). For nearly one-third (29.4%) of all antidepressant prescriptions, the drug was prescribed for an off-label indication.

Manuscript 2: In this paper, I identified the most common off-label uses for antidepressants and determined the level of scientific evidence supporting off-label antidepressant prescriptions. For this descriptive analysis, I used antidepressant prescriptions in the MOXXI system that were written between January 1, 2003 and September 30, 2015. I assigned off-label antidepressant prescriptions to one of three mutually exclusive categories: (1) strong evidence supporting use of the prescribed drug for the respective indication, (2) no strong evidence for the prescribed drug, but strong evidence for another drug in the same class for the indication, or (3) no strong evidence supporting use of the prescribed drug or any other drugs in the same class for the indication. I found that the most common off-label use for antidepressants was the use of trazodone to treat insomnia, which accounted for over one-quarter (26.2%; 95% CI, 21.9% to 30.4%) of all off-label antidepressant prescriptions. For only 15.9% (95% CI, 13.0% to 19.3%) of off-label antidepressant prescriptions, the prescribed drug had strong scientific evidence for the respective indication. For 39.6% (95% CI, 35.7% to 43.2%) of off-label antidepressant prescriptions, the prescribed drug did not have strong evidence for the respective indication but another antidepressant in the same class did. For the remaining 44.6% (95% CI, 40.2% to 49.0%) of off-label antidepressant prescriptions, neither the prescribed drug nor any other drugs in the same class had strong evidence for the indication.

Manuscript 3: In this paper, I measured the accuracy of using diagnostic codes from administrative data to determine treatment indications for antidepressant prescriptions. For this analysis, I used antidepressant prescriptions in the MOXXI system that were written

between January 1, 2003 and December 31, 2012. I linked these patients to their medical billings data (from RAMQ) and hospitalization discharge summary data (from MED-ECHO) to obtain diagnostic codes for 13 plausible conditions where antidepressants would be used. For each of these 13 indications, I determined whether a given patient had a diagnostic code for the indication recorded around his or her prescription date and compared this result to the physician-documented treatment indication recorded for the prescription in the MOXXI system. I found that the sensitivity of administrative diagnostic codes was very poor for all 13 indications, ranging from a high of only 31.2% (95% CI, 26.8% to 35.9%) for anxiety/stress disorders to as low as 1.3% (95% CI, 0.0% to 5.2%) for sexual dysfunction. Sensitivity was notably worse among older patients and patients with more chronic comorbidities. The positive predictive value of diagnostic codes varied widely between antidepressants of different therapeutic classes, where estimates were better among antidepressants that were more likely to be prescribed for the indication. Compared to hospitalization discharge summary data, medical billings data were a better source of diagnostic codes for plausible antidepressant treatment indications, with most of these codes being recorded by the prescribing physician.

Manuscript 4: In this paper, I derived a logistic regression model to predict when antidepressants were prescribed for indications besides depression. For this analysis, I used antidepressant prescriptions in the MOXXI system that were written between January 1, 2003 and December 31, 2012. Using information from the MOXXI, RAMQ, and MED-ECHO databases, I created over 370 variables that were related to characteristics of the prescription, patient, or prescriber – all of which were considered as candidate predictors of antidepressant treatment indications. I derived the final prediction model using 3-fold cross-validation methods integrated within a forward stepwise selection procedure. I assessed the performance of the final model in a held-out portion of the dataset that was not used for training. The final model included 40 covariates and had good discrimination (*c*-statistic of 0.815; 95% CI, 0.787 to 0.847), good calibration (ratio of observed to expected events 0.986; 95% CI, 0.842 to 1.136), and performed substantially better than a model containing covariates based on diagnostic codes only. The name of the molecule prescribed was by far the strongest predictor of whether

an antidepressant was prescribed for depression. Other important predictors included the presence or absence of certain diagnostic codes and drugs prescribed in the past year, the patient's age and education level, the physician's workload, and the prescribed dose.

Manuscript 5: In this paper, I used five popular machine-learning algorithms and an ensemble learning approach called "super learning" to predict when antidepressants were prescribed for indications besides depression. For this analysis, I employed the same dataset of prescriptions and covariates from manuscript 4. To optimize the hyperparameter values of each algorithm, I used a grid search procedure that assessed the performance of the algorithm iteratively over a grid of possible hyperparameter values. I then combined the predictions from the five machine-learning algorithms using super learning. I derived two super learner prediction functions: 1) a super learner whose algorithms were fit using the optimal hyperparameter values identified from the grid search procedure, and 2) a super learner whose algorithms were fit using the default hyperparameter values in the *SuperLearner* package. When these super learner functions were evaluated on a held-out portion of the dataset that had not been used for training, I found that the super learner using the optimal hyperparameter values outperformed the super learner using the default values by 4% (95% CI 1% to 8%) and had better discrimination (*c* statistic of 0.822; 95% CI, 0.795 to 0.847 compared to 0.817; 95% CI, 0.791 to 0.846). Among the five machine-learning algorithms, support vector machines performed best, followed by random forests.

Conclusions

This thesis provides important evidence showing the need for heightened pharmacovigilance around antidepressant use for off-label indications and increases capacity to perform these pharmacovigilance activities by addressing measurement challenges around treatment indications for antidepressants. This thesis also uncovers important factors associated with antidepressant treatment indications that may be used to inform health policies and interventions aimed at changing prescribing behaviours for antidepressants. Methodologically, this thesis demonstrates an approach for enabling researchers to predict treatment indications

for multiple indication drugs in the absence of documented treatment indications and makes important contributions towards improving practices for predicting outcomes in epidemiology.

Résumé

Contexte

L'Organisation mondiale de la santé décrit la pharmacovigilance comme «la science et les activités relatives à la détection, l'évaluation, la compréhension et la prévention des effets indésirables ou de tout autre problème lié à aux médicaments» (*traduction libre*). Un élément important de la pharmacovigilance est la surveillance de l'utilisation de médicaments pour des raisons médicales (indications) qui ne sont pas approuvées et manquent de données probantes. Les antidépresseurs sont un excellent exemple de médicaments nécessitant une pharmacovigilance accrue. Ces médicaments sont parmi les plus utilisés en Amérique du Nord et sont prescrits pour traiter non seulement la dépression, mais aussi de nombreuses autres indications, dont plusieurs sont non approuvées et non appuyées par des données probantes. Il est difficile de mener des activités de pharmacovigilance sur l'utilisation des antidépresseurs selon l'indication puisque cette indication est rarement documentée. En outre, il n'existe actuellement aucun mécanisme validé pour prédire cette information manquante. L'utilisation de systèmes de prescription électronique incluant l'indication pourraient permettre de résoudre ce problème et servir pour des activités de pharmacovigilance sur les antidépresseurs. Cependant, étant donné que de tels systèmes sont peu répandu, algorithmes validés qui peuvent prédire avec précision lorsque des antidépresseurs sont prescrits pour des indications autres que la dépression pourrait constituer des outils précieux pour permettre des activités de pharmacovigilance sur les antidépresseurs ainsi que d'autres médicaments pouvant être utilisés pour de multiples indications.

Objectifs

L'objectif général de cette thèse était d'accroître les capacités de pharmacovigilance concernant les antidépresseurs en utilisant les données d'un système unique de prescription électronique incluant l'indication pour répondre à quatre objectifs de recherche spécifiques:

1) Déterminer la prévalence des différentes indications de traitement, y compris les indications non approuvées, pour les antidépresseurs dans les soins primaires et évaluer le niveau de données probantes pour les indications non approuvées ;
2) Valider la précision des codes diagnostics provenant des données clinico-administratives par rapport aux indications de traitement de référence pour les antidépresseurs;
3) Utiliser des techniques standards de régression logistique pour prédire et identifier les déterminants importants des ordonnances d'antidépresseurs pour les indications autres que la dépression; et
4) Explorez l'utilisation d'algorithmes d'apprentissage machine plus flexibles et une approche d'apprentissage par ensemble appelée «super apprentissage» pour prédire quand les ordonnances d'antidépresseurs sont rédigées pour des indications autres que la dépression.

Sources de données

Ces quatre objectifs de recherche ont été atteints via cinq manuscrits. Les deux premiers manuscrits ont traité le premier objectif, tandis que les trois derniers manuscrits ont traité les trois derniers objectifs. Pour tous ces manuscrits, la source des données était MOXXI (Medical Office of the XXIst century) - un système électronique de prescription et de gestion de médicaments utilisé par environ 185 médecins de famille au Québec depuis 2003. Pour tous les patients ayant reçu une ordonnance d'antidépresseurs via le système MOXXI, des données supplémentaires sur leur utilisation des services de santé au cours de l'année précédente ont été obtenues en liant ces patients à l'ensemble des données clinico-administrative de la Régie de l'assurance maladie du Québec (RAMQ) et de gestion des données hospitalières (MED-ECHO).

Méthodes et résultats

Manuscrit 1: Dans cet article descriptif, j'ai mesuré la prévalence des différentes indications des ordonnances d' antidépresseurs qui ont été rédigées par des médecins de famille au Québec entre le 1er janvier 2006 et le 30 septembre 2015. J'ai classé les indications de traitement,

d'abord selon la condition clinique, puis selon le statut de l'approbation (approuvée vs non approuvée). Je mesurais également les changements au fil du temps dans les ordonnances d'antidépresseurs pour la dépression en utilisant des modèles de différence de risque linéaire généralisés pour estimer l'effet linéaire du temps (en années) sur la probabilité que les ordonnances d'antidépresseurs soient rédigées pour la dépression. Globalement, j'ai constaté que seulement 55,2% de toutes les ordonnances d'antidépresseurs rédigées pendant la période de l'étude étaient pour la dépression. Les médecins ont prescrit des antidépresseurs pour les troubles anxieux (18,5%), l'insomnie (10,2%), la douleur (6,1%) et les troubles paniques (4,1%). Entre 2006 et 2015, la proportion des ordonnances d'antidépresseurs rédigées pour la dépression a diminué de manière significative pour toutes les classes principales d'antidépresseurs, en particulier les inhibiteurs de la recapture de la sérotonine-norepinephrine (différence de risque à cinq ans de -9,73, IC à 95%, -11,86% à -7,61%). Près d'un tiers (29,4%) de toutes les ordonnances d'antidépresseurs étaient rédigées pour une indication non approuvée.

Manuscrit 2: Dans cet article, j'ai identifié les indications non approuvées les plus fréquentes pour les antidépresseurs, et déterminé le niveau de données probantes pour chacune de ces indications. Pour cette analyse descriptive, j'ai utilisé les ordonnances d'antidépresseurs qui avaient été rédigées dans le système MOXXI entre le 1er janvier 2003 et le 30 septembre 2012. J'ai attribué aux ordonnances d'antidépresseurs rédigées pour des indications non approuvées l'un des trois statuts suivants (mutuellement exclusifs): (1) médicament prescrit pour une indication non approuvée, mais pour laquelle des données probantes soutiennent l'usage, (2) médicament prescrit pour une indication non approuvée et non soutenue par des données probantes, mais pour laquelle un autre médicament de la même classe est soit approuvé, soit dispose de données probantes qui soutiennent l'usage, ou (3) aucune données probantes qui soutient l'usage du médicament prescrit ou de tout autre médicament de la même classe pour cette indication. J'ai constaté que l'usage non approuvé le plus courant des antidépresseurs était la trazodone pour traiter l'insomnie, qui représentait plus du quart (26,2%, IC à 95%, 21,9% à 30,4%) de toutes les ordonnances d'antidépresseurs pour des indications non approuvées. Pour seulement 15,9% (IC à 95%, 13,0% à 19,3%) des ordonnances

d'antidépresseurs hors approbation, l'usage était soutenu par des données probantes (statut 1). Pour 39,6% (IC à 95%, 35,7% à 43,2%) des ordonnances d'antidépresseurs hors approbation, l'usage du médicament n'était pas supporté par des données probantes, mais l'était pour un autre médicament de la même classe (statut 2). Pour 44,6% (IC à 95%, 40,2% à 49,0%) des ordonnances d'antidépresseurs hors approbation, ni le médicament prescrit ni aucun autre médicament dans la même classe n'était soutenu par des données probantes (statut 3).

Manuscrit 3: Dans cet article, j'ai mesuré à quelle précision les codes diagnostics issus des données administratives permettait de déterminer les indications de traitement des ordonnances d'antidépresseurs. Pour cette analyse, j'ai utilisé les ordonnances d'antidépresseurs dans le système MOXXI qui ont été rédigées entre le 1er janvier 2003 et le 31 décembre 2012. J'ai relié les patients à leurs données de facturation médicale (de la RAMQ) et d'hospitalisation (de MED-ECHO) et obtenu les codes de diagnostics de 13 conditions pour lesquelles il est plausible que les antidépresseurs soient utilisées. J'ai déterminé si un patient donné avait un code diagnostic enregistré pour une de ces 13 indications autour de la date d'ordonnance d'antidépresseurs, et comparé ce code à l'indication documentée par le médecin sur l'ordonnance dans le système MOXXI. J'ai constaté que la sensibilité des codes diagnostics en provenance des données administratives était très faible pour les 13 indications, allant d'un maximum de seulement 31,2% (IC à 95%, 26,8% à 35,9%) pour les troubles de l'anxiété et du stress à 1,3% (95% CI, 0,0% à 5,2%) pour un dysfonctionnement sexuel. La sensibilité était moins bonne notamment pour les patients plus âgés et les patients avec plus de comorbidités chroniques. La valeur prédictive positive des codes diagnostics variait largement entre les antidépresseurs de différentes classes thérapeutiques, et les estimations étaient meilleures lorsque les antidépresseurs étaient plus susceptible d'être prescrits pour une indication. Par rapport aux données d'hospitalisation, les données sur les facturations médicales étaient une meilleure source de codes diagnostic pour les indications plausibles du traitement antidépresseur, la plupart de ces codes étant enregistrés par le médecin prescripteur dans les données de facturation.

Manuscrit 4: Dans cet article, j'ai dérivé un modèle de régression logistique pour prédire quand les antidépresseurs étaient prescrits pour des indications autres que la dépression. Pour cette analyse, j'ai utilisé des ordonnances d'antidépresseurs dans le système MOXXI qui ont été rédigées entre le 1er janvier 2003 et le 31 décembre 2012. En utilisant les informations provenant des bases de données MOXXI, RAMQ et MED-ECHO, j'ai créé plus de 370 variables liées aux caractéristiques de l'ordonnance, du patient ou du prescripteur - qui ont toutes été considérées comme des prédicteurs candidats de l'indication de traitement de l'antidépresseur. J'ai dérivé le modèle de prédiction définitif en utilisant des méthodes de validation croisée 3 fois, intégrées dans une procédure de sélection par étape. J'ai évalué la performance du modèle final dans une portion des données qui n'avait pas été utilisée pour l'entraînement du modèle. Le modèle final incluait 40 covariables et avait une bonne discrimination *(c* -statistic de 0,815; IC à 95%, 0,787 à 0,847), une bonne calibration (rapport des taux observés aux événements attendus 0,986; IC à 95%, de 0,842 à 1,136), et était nettement plus performant qu'un modèle contenant uniquement les covariables basées sur les codes diagnostic. Le nom de la molécule prescrite était de loin le prédicteur le plus puissant pour déterminer si un antidépresseur était prescrit pour la dépression. D'autres prédicteurs importants comprenaient la présence ou l'absence de certains codes diagnostic et certains médicaments prescrits au cours de l'année précédente, l'âge et le niveau de scolarité du patient, la charge de travail du médecin et la dose prescrite.

Manuscrit 5: Dans cet article, j'ai utilisé cinq algorithmes d'apprentissage machine courants et un ensemble d'approches d'apprentissage appelé "super apprentissage " pour prédire quand les antidépresseurs sont prescrits pour des indications autres que la dépression. Pour cette analyse, j'ai utilisé le même ensemble de données d'ordonnances et covariables que dans le manuscrit 4. Afin d'optimiser les valeurs des hyperparamètres de chaque algorithme, j'ai utilisé une procédure de recherche structurée qui a évalué de manière itérative la performance de l'algorithme sur une grille de valeurs possibles des hyperparamètres. Ensuite, j'ai combiné les prédictions des cinq algorithmes d'apprentissage machine utilisant le "super apprentissage". J'ai tiré deux super fonctions de prédiction de cet apprentissage: 1) un "super –

apprenant" dont les algorithmes ont été adaptés en utilisant les valeurs des hyperparamètres optimales identifiées à partir de la procédure de recherche structurée, et 2) un "super apprenant" dont les algorithmes ont été adaptés en utilisant les valeurs des hyperparamètres par défaut dans le package *SuperLearner.* Lorsque ces fonctions super apprenant ont été évaluées sur une partie des données qui n'avait pas été utilisée pour la formation, j'ai trouvé que le "super apprenant" qui utilisait les valeurs optimales des hyperparamètres a mieux performé que le "super apprenant" qui utilisait les valeurs par défaut, à hauteur de 4% (IC à 95% 1% à 8%) et avait une meilleure discrimination *(c* statistique de 0,822, IC à 95%, de 0,795 à 0,847 par rapport à 0,817; 95% CI, de 0,791 à 0,846). Parmi les cinq algorithmes d'apprentissage machine, les machines à vecteurs de support ont été les meilleurs, suivis des forêts d'arbres décisionnels.

Conclusions

Cette thèse présente des données importantes montrant la nécessité d'une pharmacovigilance accrue autour de l'utilisation des antidépresseurs pour des indications non approuvées, et augmente la capacité d'effectuer ces activités de pharmacovigilance en abordant les défis de mesure autour des indications de traitement des antidépresseurs. Cette thèse révèle également des facteurs importants associés aux indications de traitement antidépresseur, qui peuvent être utilisés pour informer les politiques de santé et les interventions visant à modifier les comportements de prescription des antidépresseurs. Méthodologiquement, cette thèse démontre une approche permettant aux chercheurs de prédire les indications de traitement pour les médicaments à indications multiples en l'absence d'indications de traitement documentées, et apporte d'importantes contributions à l'amélioration des pratiques de prévision des résultats en épidémiologie.

Acknowledgements

It is hard to believe that I am finally at the end of my PhD journey. These past five years have been the most challenging, yet rewarding years of my academic life. I thank God for giving me the opportunity to come to McGill and purse my PhD, and I feel very fortunate to have worked with so many great people. I would like to take a moment to express my gratitude to all the individuals who helped me along this journey.

First and foremost, I would like to acknowledge my supervisor, Dr. Robyn Tamblyn. Thank you for giving me the opportunity to pursue a unique research idea that eventually blossomed into a research interest of my own. Beyond this, you provided me with the perfect balance between academic support and freedom that allowed me to take ownership of my work and become confident in my research skills. Thank you for always encouraging me to strive for the best and take on challenges. I have learned so much from you and am eternally grateful to have had the opportunity to work with you.

I would also like to thank my thesis committee members, Drs. Michal Abrahamowicz and David Buckeridge. Dr. Abrahamowicz – thank you for always being the first to get back to me with such excellent comments and feedback. Dr. Buckeridge – thank you for offering your insights on my work and valuable advice regarding my future career plans.

A special thanks to Dr. Aude Motulsky who took an interest in my thesis and graciously provided me with valuable clinical input on my work. The first two manuscripts of this thesis would not have been possible without your clinical expertise. Thank you for tirelessly reviewing draft after draft of my manuscripts, meeting with me countless times to refine our message, answering multiple emails from me daily, and even helping me translate my thesis abstract. It has been an absolute pleasure working with you and I will miss our collaborations.

Thank you to all my classmates at McGill for your support and encouragement throughout the years. I could not have asked for a better PhD cohort. A special thanks to Ania, Daniala, Jenn,

and Kathryn who made this journey a more pleasant experience. I am thankful for your friendship, and I hope we can stay in touch even after we go our separate ways.

I would also like to thank the administrative staff at McGill for all their help over the years. Special thanks go out to Patricia Plouffe, Rosalba Pupo, Katherine Hayden, and Andre Yves Gagnon for always being willing to help me book meetings, answer my questions, deal with my funding issues and travel expenses, secure examiners, etc.

To Travis – I am so thankful for a boyfriend who understands in *every* way what this experience has been like. Thank you for always believing in me and encouraging me to take the harder yet more rewarding route. I would not have attempted some of the machine-learning analyses in this thesis if you had not encouraged me to tackle my programming fears. Thank you for your companionship through many afternoons, late nights, weekends, and even holidays as we worked on our PhDs together. I will miss working beside you! I appreciate all your prayers, programming help, and feedback on my thesis work. You know my PhD so well that you can explain it to our friends even better than I can. When it comes time for you to finish up, I will be there to cheer you on, just as you did for me.

And last, but certainly not least, a big thank you to my parents Rose and Daniel, sister Serena, and brother-in-law Mac for all their prayers and encouragement from day one of my PhD. To my mom and dad, thank you for being there for me through every step of this journey – the good and the bad, the breakthroughs and the breakdowns. Your daily encouraging texts have kept me going, and even though you live in Ottawa, you still try to help me out by coming here on the weekends to stock my fridge with groceries and frozen meals. Thank you for not minding that I missed so many events at home because I was busy working on my thesis. I am eternally grateful for you both and would not be where I am today without your love and support.

Contributions of authors

The findings of this thesis were reported in five manuscripts:

Manuscript 1: 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.

Manuscript 2: 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. *The BMJ*. 2017;356:j603.

Manuscript 3: Wong J, Abrahamowicz M, Buckeridge DL, Tamblyn R. Assessing the accuracy of using diagnostic codes from administrative data to infer antidepressant treatment indications: a validation study.

Manuscript 4: Wong J, Abrahamowicz M, Buckeridge DL, Tamblyn R. Derivation and validation of a model to identify when antidepressants are prescribed for indications besides depression: a prediction study.

Manuscript 5: Wong J, Manderson T, Abrahamowicz M, Buckeridge DL, Tamblyn R. Optimizing the hyperparameter values of machine-learning algorithms improved the performance of super learning for predicting when antidepressants were prescribed for indications besides depression.

The objectives of these manuscripts were developed in collaboration with my supervisory committee. In all manuscripts, I defined the specific research question, reviewed the necessary methodological and substantive literature, developed the study design, created the study datasets, performed all data management and statistical analyses, interpreted the findings, and wrote the first draft of each manuscript. All co-authors have reviewed and approved the final version of each manuscript.

Robyn Tamblyn is a James McGill Chair, a Professor in the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University, the Scientific Director of the Institute of Health Services and Policy Research at the Canadian Institutes of Health Research, and Scientist at the Research Institute of the McGill University Health Centre. As my thesis supervisor, Dr. Tamblyn provided me with access to all the necessary data, provided support throughout all stages of the manuscript development process, and provided feedback on all five manuscripts and chapters in this thesis.

Michal Abrahamowicz is a James McGill Professor in the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University. As a member of my supervisory committee, Dr. Abrahamowicz provided support regarding the statistical methodology and interpretation of results in all manuscripts. He also provided valuable feedback on all five manuscripts, particularly on the sections describing the statistical methods.

David L. Buckeridge is an Associate Professor in the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University and holds a Canada Research Chair in Public Health Informatics. As one of my thesis committee members, Dr. Buckeridge provided input on the interpretation of the results and reviewed all five manuscripts to provide substantive feedback.

Aude Motulsky is a Researcher-Professor in the Department of Management, Evaluation and Health Policy in the School of Public Health at the Université de Montréal. Dr. Motulsky also holds a BPharm and is a practicing pharmacist. Dr. Motulsky provided extensive input on the design and interpretation of results in the first two manuscripts. She also reviewed the first two manuscripts and provided valuable clinical feedback.

Tewodros Eguale is an Associate Professor in the School of Pharmacy at the Massachusetts College of Pharmacy and Health Sciences. He was involved in the first two manuscripts and provided input on the study design and substantive feedback on the two manuscripts.

Travis Manderson is a doctoral candidate in the School of Computer Science at McGill University. He is knowledgeable and experienced in machine-learning and provided methodological and programming support for the fifth paper. He also reviewed the paper and provided editorial feedback.

Statement of originality

The work in this thesis represents original contributions that further knowledge about antidepressant treatment indications and develop new methods to measure them. Manuscript 1 is the first study to use gold-standard treatment indications from an electronic prescribing system to perform a detailed analysis of treatment indications for antidepressant prescriptions. Manuscript 2 is the first study to analyze the level of scientific evidence supporting off-label indications for antidepressant prescriptions and investigate the extent to which intra-class substitution may contribute to off-label prescribing. Manuscript 3 is the first study to validate the standard approach of using diagnostic codes from administrative data to determine antidepressant treatment indications. Because the findings from manuscript 3 showed that diagnostic codes were poor markers for antidepressant treatment indications, manuscripts 4 and 5 derived algorithms that used other variables from health services data beyond diagnostic codes to accurately predict when antidepressant were prescribed for indications besides depression. To my knowledge, manuscripts 4 and 5 are the first studies to derive algorithms that can accurately predict gold-standard treatment indications for antidepressant prescriptions.

This thesis also makes an original contribution towards improving practices for predictive modelling in epidemiology. Manuscript 5 is the first study to investigate the effect of hyperparameter tuning on the predictive performance of the popular ensemble machine-learning approach called “super learning”. The finding that the performance of super learning improved when the hyperparameter values of machine-learning algorithms were optimally tuned demonstrates an important lesson that will help investigators obtain even better results when using super learning.

Although I received guidance from my thesis supervisory committee and other co-authors on the substantive and methodological aspects of my thesis, I confirm that the conception, execution, and drafting of the work in this thesis were done entirely on my own.

Statement of financial support

I am very grateful for all the financial support that I received over the course of my doctoral studies at McGill University. I was very fortunate to receive the Vanier Canada Graduate Scholarship, which provided me with generous funding for the first three years of my doctoral studies. The fourth year of my doctoral studies was jointly funded by Dr. Tamblyn and the Max E. Binz Fellowship from the Faculty of Medicine at McGill University. In the final year of my doctoral studies, I was grateful to receive a studentship and fellowship award from the Research Institute of the McGill University Health Center that was once again supplemented by a gracious stipend from Dr. Tamblyn.

I am also thankful for all the travel funding I received, which enabled me to present my thesis work at international and national conferences. In 2015, I presented the findings from my first manuscript at the *31st International Conference on Pharmacoepidemiology and Therapeutic Risk Management* in Boston with the funding that I received from a Graduate Research Enhancement and Travel Award from the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University and an Institute Community Support travel award from the Canadian Institutes of Health Research. In 2016, I presented the findings from my third manuscript at the *2016 Epidemiology Congress of the Americas* conference in Miami with funding provided by Dr. Tamblyn. Most recently, I presented the findings from my fourth manuscript at the *2017 Canadian Association for Health Services Policy and Research* conference in Toronto with the funding I received once again from an Institute Community Support travel award from the Canadian Institutes of Health Research.

Table of contents

List of tables

List of figures

List of abbreviations

CI	Confidence interval
ICD-9	International classification of diseases, Ninth revision
ICD-10	International classification of diseases, Tenth revision
ICPC	International Classification of Primary Care
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
LR-	Negative likelihood ratio
LR+	Positive likelihood ratio
MAOI	Monoamine oxidase inhibitor
MED-ECHO	Quebec's hospital discharge summary database
MOXXI	Medical Office of the XXIst Century
NLP	Natural language processing
NPV	Negative predictive value
PPV	Positive predictive value
RAMQ	Régie de l'assurance maladie du Québec
RBF	Radial basis function
RE	Relative efficiency
SNRI	Serotonin norepinephrine reuptake inhibitor
SSRI	Selective-serotonin reuptake inhibitor
TCA	Tricyclic antidepressant
US	United States of America

1 Introduction

1.1 Context

Pharmacotherapy (the treatment of disease with drugs) plays an integral role in contemporary medicine. Prescription drugs are taken by 55% of Canadians aged 45 to 64 years and 83% of Canadians aged 65 to 79 years (1). In fact, nearly one-third (30%) of Canadians aged 65 to 79 years concurrently take five or more medications (1). Not surprisingly then, prescription drugs represent a significant source of health care spending, constituting the third most costly component of health care in Canada (2). In 2013, prescription drugs accounted for nearly 14% ($29.3 billion) of Canada's total health spending (3).

Although drugs have the potential to prevent and treat health problems, they can also produce harm. Even when drugs are properly administered in the correct dose, they can still cause adverse effects, which are referred to as adverse drug reactions (ADRs) (4). In 2010-2011, more than 27,000 seniors in Canada were hospitalized for ADRs – increasing at an average rate of 3% per year since 2006-2007 (5). Thus, besides being noxious to patients, ADRs are also expensive, costing an estimated $35.7 million due to ADR-related hospitalizations and emergency department visits among elderly patients in Canada (6).

A major cause of ADRs is the inappropriate use of medications, which can arise from drug-drug interactions or drug use that is either off-label or contraindicated (7). The ability to detect and reduce inappropriate medication use is important because such inappropriate uses can lead to ADRs that are potentially preventable (7).

1.2 Pharmacovigilance and off-label drug use

The World Health Organization defines pharmacovigilance as "the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems" (8). A very relevant issue to the science of pharmacovigilance is the use of medications for indications that are not approved ("off-label") and for which inadequate

evidence exists to support their use (8). Indeed, this issue is a concern as an estimated 11% to 21% of all prescriptions are for off-label indications, and approximately three-quarters of these off-label prescriptions do not have strong evidence supporting the drug's use for the indication (9,10). Moreover, the risk of ADRs has been found to be 54% higher when drugs are used for off-label indications without strong evidence than when drugs are used for on-label indications (9).

Pharmacovigilance efforts aimed at monitoring or evaluating drug use for off-label indications absolutely require data on the medical reasons (i.e. treatment indications) for prescriptions. Without this information, patients who have been prescribed drugs for off-label indications cannot be tracked or followed for safety and effectiveness outcomes, which could otherwise improve prescribing practices for medications and inform future labeling revisions (11).

1.3 Antidepressants and the need for increased pharmacovigilance

Antidepressants are one of the most commonly used medications in Canada. Between 1981 and 2000, the total number of dispensed antidepressant prescriptions increased by 353% from 3.2 million to 14.5 million (12). Currently, antidepressants are the top prescription medication used among Canadian women, with nearly one in seven women aged 25 to 79 taking these drugs (1). Antidepressants are also the top medication used among Canadian males aged 25 to 44 (1). In fact, according to a study by the Organization for Economic Co-operation and Development (OECD), Canada has the third highest per-capita consumption level of antidepressants out of 23 OECD countries (13).

An important factor driving the widespread use of antidepressants is a broadening of indications for these drugs (14–16). Antidepressants are reportedly used to treat a host of indications besides depression including generalized anxiety disorders, social phobia, obsessive-compulsive disorder, panic disorders, sleeping disorders, neuropathic pain, urinary incontinence, migraines, eating disorders, and even premature ejaculation (15,17). While some of these indications are approved for antidepressants, others are not approved and even lack

supporting scientific evidence (17). Thus, the widening of indications for antidepressants not only complicates an evaluation of the appropriate use of antidepressants, but also raises pharmacovigilance concerns around inappropriate drug use (18).

For example, the antidepressant trazodone is one of the most commonly used pharmacotherapies for insomnia (19,20). However, numerous review articles (20–22) and clinical guidelines (19,23) warn that minimal scientific evidence exists to support trazodone's efficacy in treating insomnia, particularly among patients without depression. The potentially ineffective use of trazodone to treat insomnia calls for increased pharmacovigilance because trazodone can cause notable side effects including nausea, vomiting, cardiovascular complications, drowsiness, dizziness, and priapism – a rare but serious side effect that can occur even at the low doses used for insomnia (19,21). Besides trazodone, non-evidence-based off-label use of newer generation antidepressants like selective serotonin reuptake inhibitors (SSRIs) and serotonin noradrenaline reuptake inhibitors (SNRIs) is also a concern because although these drugs are considered safer and more tolerable than the older tricyclic antidepressants (TCAs), they can still cause serious and bothersome adverse side effects including gastrointestinal symptoms, hepatotoxicity, cardiovascular disturbances, osteoporosis, sexual dysfunction, and weight gain (24,25).

Given these safety concerns and the potential that much of the widespread use of antidepressants may be for non-evidence-based indications, efforts to increase pharmacovigilance around antidepressants are urgently needed. Examples of valuable pharmacovigilance activities include monitoring the extent to which antidepressants are being prescribed for different treatment indications, identifying priority non-evidence-based off-label antidepressant uses (i.e. drug-indication pairs) in need of further evaluation, and conducting studies to assess the safety and effectiveness of priority off-label uses. To carry out these activities, however, it is essential to first know the treatment indication for a given antidepressant prescription. Without this information, antidepressant use for off-label indications cannot be tracked, let alone evaluated. In fact, studies evaluating antidepressant

use for *any* indication – on-label or off-label – are susceptible to bias if treatment indications are not accounted for. For example, studies measuring the frequency of pharmacotherapy for depression over time may be biased if antidepressant prescriptions are used as a proxy for depression treatment, particularly if the proportion of antidepressants being prescribed for indications besides depression is substantial and/or changing over time. Effectiveness studies could also be biased if the outcomes of interest are not relevant to all antidepressant users (e.g. chronic pain management). Finally, studies comparing the safety of different antidepressants could be confounded if certain antidepressants are more likely to be prescribed for treatment indications involving patient characteristics that are also associated with the outcome (i.e. confounding by indication). For example, myocardial infarction is a safety outcome that has been studied for antidepressants (26). However, the antidepressant bupropion is also approved to treat nicotine dependence among smokers (27), and smokers are more likely to experience myocardial infarctions (28). Thus, if treatment indications are not accounted for in the analysis, one may observe a spurious association between bupropion use and myocardial infarction that is simply because a higher proportion of bupropion users are heavy smokers compared to other antidepressant users.

1.4 Measuring treatment indications for antidepressant prescriptions

Given the vital role that treatment indication data plays in enabling pharmacovigilance activities for antidepressants, it is important to have sound methods for measuring antidepressant treatment indications. However, treatment indications for antidepressants are challenging to measure for several reasons. First, health insurance agencies do not require information on treatment indications for prescriptions to reimburse medications, thus this information is not recorded in prescription claims databases. Second, pharmacies do not need to know the treatment indication for medications before dispensing them, thus this information is usually not recorded in pharmacy databases either. Finally, even in the medical chart, many physicians do not explicitly state a direct link between medications and their corresponding indications, which can create uncertainty about the correct drug-indication pair even if one has access to the problem list (29).

Despite these challenges, several studies (30–34) have attempted to infer antidepressant treatment indications by searching for the presence of diagnostic codes for on-label and well-known off-label indications recorded for patients in administrative health databases within a pre-specified time window. Although this approach is simple and feasible to apply to large health databases, it has several limitations. First, diagnostic codes are often incomplete and inaccurate, especially for mental health conditions (35). In fact, studies that have employed diagnostic codes could not assign a treatment indication to between 25% (32) and 50% (31) of antidepressant users because a diagnosis of interest could not be found. Second, among antidepressant users with a diagnosis of interest, studies have found that up to a quarter of these users had diagnostic codes for two or more indications (31), creating uncertainty about the main or most responsible indication. Finally, because the accuracy of this approach of using diagnostic codes to infer antidepressant treatment indications has never been validated against a gold standard, the validity of conclusions from such studies employing this approach is unknown.

Taken together, the absence of treatment indication data for medications in health services databases and the limitations of current approaches for predicting antidepressant treatment indications create significant barriers to conducting pharmacovigilance activities for antidepressants.

1.5 The Medical Office of the XXIst Century (MOXXI): an opportunity to conduct pharmacovigilance activities for antidepressants

The Medical Office of the XXIst Century (MOXXI) is an electronic prescribing and drug management system that was developed by researchers at McGill University to enhance patient safety in primary care (36). The MOXXI system addresses specific functionalities needed to improve the safety and quality of drug management, including (a) enabling physicians to write electronic prescriptions that are transmitted to community pharmacy systems, (b) retrieving and displaying all drugs dispensed to patients in the past 12 months from community

pharmacy systems, (c) generating automated alerts about potential prescribing problems (e.g. drug-disease, drug-drug, drug-age, and drug-allergy contraindications) from a drug knowledge base that consults all documented allergies, diseases, and actively prescribed and dispensed medications for patients, and (d) allowing physicians to make modifications (e.g. stop or change orders) to existing prescriptions that are then transmitted to the original dispensing pharmacy (36). The MOXXI system also allows physicians to access other clinical data on patients besides dispensed drugs, including their health problem list, allergy list, and recent hospitalizations and emergency department visits. These features are possible because the MOXXI system is integrated with administrative datasets from the provincial health insurance agency (The Régie de l'assurance maladie du Québec [RAMQ]) and the provincial hospital discharge summary database (MED-ECHO). These data sources provide information on patient demographics, dispensed medications, diagnoses, hospitalizations, and medical services received.

In addition to these functionalities, an important and unique feature of the MOXXI system is that physicians are required to document at least one treatment indication for each drug prescribed by either selecting from a drop-down list of on-label and off-label indications (without distinction) or by typing the indication(s) into a free-text field (Figure 1-1). Physicians have incentive to record the true indication for the prescription (rather than select the first indication listed in the drop-down menu, for example) because these indications become part of the patient's problem list, which are then used by the drug knowledge base to identify potential drug-disease contraindications and generate pop-up alerts (36). In fact, the treatment indications in the MOXXI system were previously validated again a blinded, post-hoc physician-facilitated chart review where they had excellent sensitivity (98.5%) and high positive predictive value (97.0%) (37).

The presence of gold-standard treatment indications in the MOXXI system linked to patient-level health services data from RAMQ creates an unprecedented opportunity to conduct important pharmacovigilance activities for antidepressants. First, the MOXXI indications could be used to measure the prevalence of different treatment indications for antidepressants,

including off-label prescribing. Second, diagnostic codes for antidepressant treatment indications in the RAMQ data could be validated against the gold-standard MOXXI indications to measure the accuracy of using administrative diagnostic codes to predict antidepressant treatment indications. Finally, to improve the ability to predict antidepressant treatment indications, predictive models for gold-standard MOXXI indications could be derived using variables in RAMQ beyond diagnostic codes. Such models would not only uncover important factors associated with antidepressant treatment indications, but would also allow researchers in other settings to impute missing treatment indications, thus enabling research on antidepressant use by indication even in the absence of gold-standard treatment indications.

1.6 Modern approaches for predicting complex outcomes like antidepressant treatment indications

The paucity of research on antidepressant treatment indications creates challenges for optimally predicting them, particularly because little is known about determinants of this outcome. One solution is to consider all the information available and employ a data-mining approach to identify important predictors. However, there may be many variables to consider and the true associations may be complex to model. For example, the data may contain intricate non-linear associations that are not well represented by individual covariate terms or there may be higher-order interactions that exist between variables. These complexities are challenging to capture using standard parametric regression models (e.g. logistic regression) because these models must be fully and correctly specified to achieve unbiased probability estimates (38). Thus, as an alternative to standard regression models, it may be worthwhile to explore the use of more flexible, data-driven (i.e. non-parametric) prediction methods from the machine-learning literature to predict antidepressant treatment indications. However, because there are many algorithms to choose from and investigators may not know which algorithm to use a priori, many investigators have turned to the ensemble machine-learning approach called "super learning" for their prediction tasks (39). With super learning, investigators do not have to choose just one algorithm. Instead, they can choose a collection of algorithms and combine their predictions into one "super learner" prediction function. This method is quickly becoming

a popular approach for predicting outcomes in epidemiology, and investigators can easily implement this method using the *SuperLearner* package (40) in the R programming language. However, it appears that when investigators use *SuperLearner*, they often run the machine-learning algorithms "as is" – using the default hyperparameter values in the package (41–45). However, because these hyperparameters control the complexity of the algorithms, their values should be chosen carefully to optimize algorithm performance (46). Thus, compared to the default values, optimizing the hyperparameter values of machine learning algorithms may improve the performance of super learning for predicting antidepressant treatment indications.

1.7 Research objectives

The overall goal of this thesis was to leverage the unique data in the MOXXI system to perform activities to increase the capacity for comprehensive pharmacovigilance around antidepressants. To achieve this goal, I addressed four specific research objectives:

1) Determine the prevalence of different treatment indications, including off-label indications, for antidepressants in primary care and to assess the level of scientific support for off-label indications;
2) Evaluate the accuracy of using diagnostic codes recorded in administrative health data to infer treatment indications, compared to gold-standard treatment indications for antidepressants;
3) Use standard regression techniques to predict and identify important determinants of antidepressant prescriptions for indications besides depression; and
4) Explore the use of more flexible machine-learning algorithms and the ensemble learning approach called "super learning" to predict when antidepressant prescriptions are written for indications besides depression.

1.8 Organization of the thesis

This thesis is organized around five research manuscripts. Chapter 2 provides further background for the manuscripts by reviewing the methods and findings from previous studies that have attempted to measure treatment indications for antidepressants. Chapter 3

describes the data sources and criteria used to create the datasets in the five manuscripts. The next five chapters (Chapters 4 to 8) contain the manuscripts that address the thesis objectives. Chapter 4 describes the prevalence of different treatment indications for antidepressants according to their clinical conditions and their label status, and measures trends in antidepressant prescribing for depression versus other indications over the past decade. This manuscript was published in the *Journal of the American Medical Association* under the title, "Treatment indications for Antidepressants Prescribed in Primary Care in Quebec, Canada, 2006-2015." Chapter 5 takes a closer look at off-label indications for antidepressants and assesses the level of scientific evidence supporting different off-label antidepressant uses. This manuscript was published in *The BMJ* under the title, "Off-label indications for antidepressants in primary care: descriptive study of prescriptions from an indication based electronic prescribing system." Chapter 6 reports the findings of a validation study assessing the accuracy of administrative diagnostic codes for plausible antidepressant treatment indications compared to gold-standard treatment indications for antidepressants recorded in the MOXXI system. This manuscript will be submitted to *Epidemiology* under the title, "Assessing the accuracy of using diagnostic codes from administrative data to determine antidepressant treatment indications: a validation study." Chapters 7 and 8 both determine how well variables from health services data can predict when antidepressants are prescribed for indications besides depression. Chapter 7 describes the details of a prediction model derived using classical regression techniques, the findings of which will be submitted to *The BMJ* under the title, "Derivation and validation of a model to identify when antidepressants are prescribed for indications besides depression: a prediction study." Chapter 8 explores the use of more flexible machine-learning algorithms for this same prediction task and determines whether optimizing their hyperparameter values of these algorithms improves the performance of super learning. This manuscript will be submitted to the *International Journal of Epidemiology* under the title, "Optimizing the hyperparameters of machine-learning algorithms improved the performance of super learning for predicting when antidepressants were prescribed for indications besides depression." Finally, Chapter 9 summarizes the main findings and contributions of this thesis,

highlights some important considerations, and discusses new directions for future research on antidepressants.

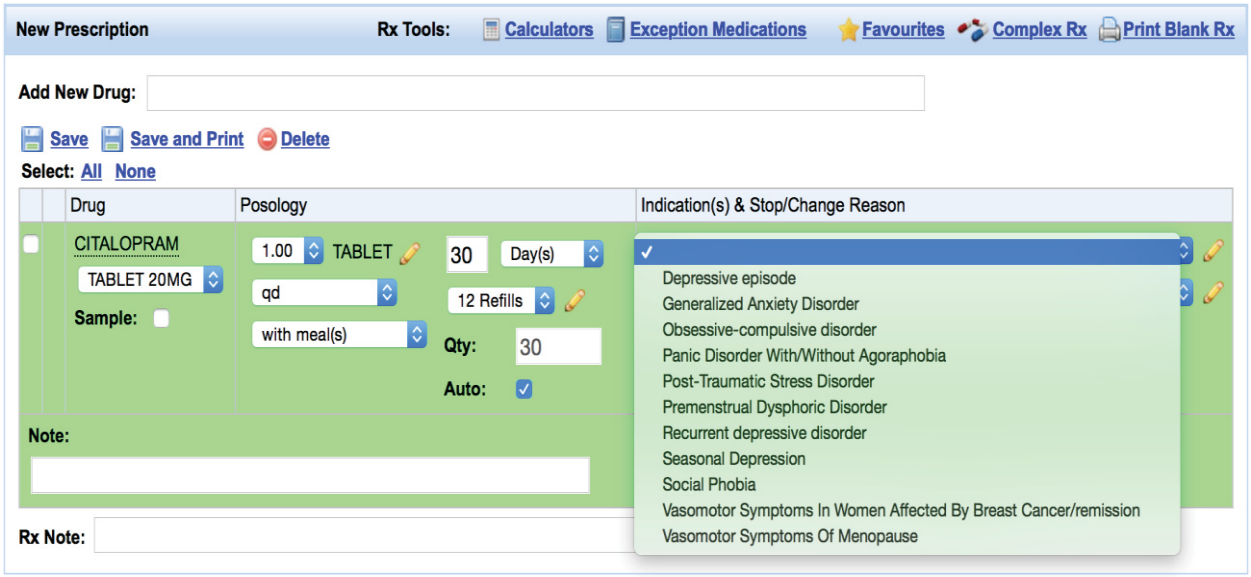


Figure 1-1. Documentation of treatment indications in the MOXXI system

2 Background

In this chapter, I conduct a literature review of published studies that have measured treatment indications for antidepressant prescriptions in any part of the analysis. I summarize the characteristics of these studies, the measurement methods they use, and their main findings. Finally, I discuss the limitations of these studies and highlight some advantages of using indication-based electronic prescribing systems to measure antidepressant treatment indications.

2.1 Previous studies that have measured treatment indications for antidepressants

I searched the PubMed database for relevant studies using two search strings: (1) antidepressant* AND indication* AND prescriptions, and (2) antidepressant* AND reason* AND prescriptions. As of May 12, 2017, these search strings returned 135 and 108 eligible studies, respectively. After reviewing the titles and abstracts of these studies, I retrieved 41 studies after removing irrelevant articles, duplicate articles, and articles not written in English. I obtained the full manuscripts of these 41 studies and extracted the following information: (1) author names, (2) study population/data source, (3) period over which the data were collected, (4) antidepressants included, (5) methodology for measuring treatment indications, and (6) main findings. Table 2-1 shows the information that was extracted from each study.

2.1.1 Study characteristics

Most of the 41 studies were from Europe (16–18,47–69) or the USA (14,31,33,34,70–76), but there were also studies from Canada (9,77), Australia (78), and Taiwan (32). Most studies were conducted quite recently as 33 (9,14,17,18,31,32,34,51,53,56,58,59,62,63,66–69,71–75,77) of them used data that were collected in 2000 or later. Thirteen (9,17,31–34,52,53,64,66,69,72,74) studies had large patient populations of over 10,000 patients while five studies (47,48,51,63,78) had very small patient populations of less than 100 patients. Eight studies (47,56,58,60,63–65,75) included children and/or adolescents only, four studies (51,53,57,68) included elderly patients only, while the remaining studies included patients of all ages. Most studies included all major classes of antidepressants (i.e. SSRIs, SNRIs, and TCAs) and

atypical antidepressants (e.g. bupropion, trazodone, mirtazapine), but eight (48–50,52,55,69,70,78) studies included only a subset of antidepressants. In seven studies (31,32,50,61,64,65,71), the analysis only considered new antidepressant prescriptions where the length of time during which patients could not have used antidepressants prior to the index date was usually one year; however, one study only required six months (71) while another study required five years (50).

2.1.2 Methods used to measure treatment indications for antidepressants

Among the 41 studies, there were a variety of different methods used to measure antidepressant treatment indications. Only four of these studies (9,49,56,78) used data from a medical record system that linked medications *directly* to their indications. The first study (49) used data from a problem-oriented medical record system in northern Sweden where physicians at a multidoctor district health centre instructed their secretaries to document all prescribed medications along with their corresponding diagnoses. However, only 12 physicians worked at this health centre and only 257 prescriptions for tricyclic antidepressants written in 1973 were included in the analysis. The second study (78) used information from a computer-based monitoring system in Sydney, Australia that recorded all drugs prescribed by general practitioners linked to their appropriate problems. However, this system was only used by three physicians at one general practice and the results included only 59 prescriptions for amitriptyline written in 1981. The third study (56) used data from the Netherlands Information Network of General Practice – a nationally representative database containing information on prescribed medications and their corresponding physician-diagnosed indications that were recorded using International Classification of Primary Care (ICPC) codes. This study included all antidepressant prescriptions that were written for children under 18 years of age in 97 general practices in 2001 and 73 general practices in 2005. However, ICPC codes were recorded for only 58% and 61% of the antidepressant prescriptions issued in 2001 (n_{total}=732) and 2005 (n_{total}=458), respectively. Finally, the fourth study (9) used data from the MOXXI electronic prescribing and drug management system (introduced in Chapter 1) that required physicians to document treatment indications for all medications at the time of prescribing. The analysis

included over 16,000 antidepressant prescriptions written by primary care physicians in Quebec between 2005 and 2009. However, because the study did not focus specifically on antidepressants, it only reported the overall proportion of off-label indications for antidepressant prescriptions. The off-label prescribing rates for individual antidepressants were not described, nor was the frequency of different clinical indications for antidepressants.

For the remaining 37 studies, 15 studies *indirectly* linked antidepressant prescriptions to treatment indications using diagnostic codes recorded in administrative (30–34) or electronic health record databases (17,53,59–62,64–66,69); 16 studies (14,48,52,55,57,58,63,67,68,70,71,73–77) contacted physicians via questionnaires, interviews, or surveys to determine treatment indications; three studies (50,51,54) examined patients' clinical notes for mentions of treatment indications, two studies (16,18) conducted structured interviews with antidepressant users for self-reported treatment indications, and one study (47) attempted to identify treatment indications by examining antidepressant prescriptions. Among the 16 studies that contacted physicians, seven (14,70,73–77) studies used a survey design whereby a nationally representative sample of physicians recorded all drugs recommended to patients ("drug mentions") and the medical reasons for these recommendations. Physicians reported this information periodically (e.g. two consecutive days per quarter) in diaries or confidential logbooks.

2.1.3 Main findings

Twenty-one studies (14,16–18,30,33,49–51,53,58,62,64,66–68,71,74,76–78) found that approximately 40% to 60% of antidepressant prescriptions were for depression. Two studies (31,32) reported notably lower percentages (20% to 30%) – most likely because they used diagnostic codes from administrative data to measure treatment indications, which are often incomplete for mental health conditions (35). On the other hand, three studies (52,57,70) found that over 70% of antidepressant prescriptions were for depression – all of which used either physician surveys (70), physician questionnaires (52), or physician documentation on study forms (57) to measure treatment indications. In one of these studies (52), the high

prevalence of treatment indications for depression (71% to 83%) was likely because the analysis only included prescriptions for four SSRIs that were issued within a year after the respective SSRI was released on the market – a time when these drugs would have been most likely prescribed for depression. In the second study (70), the high prevalence (78%) of treatment indications for depression may have been because the analysis included prescriptions written by primary care physicians *and* psychiatrists, whereas in other studies, the analysis included prescriptions written by either primary care physicians only or physicians of all specialties. In the third study (57), the fact that only nursing home residents were included in the analysis may help explain the high prevalence (77%) of treatment indications for depression, particularly since depression is a common mental illness among nursing home residents (79).

In terms of trends over time, four studies (14,55,76,77) found that the proportion of antidepressants being prescribed for depression had declined over time. All these studies measured treatment indications using physician survey data (14,76,77) or physician documentation (55). O'Brien *et al.* (76) found that percentage of antidepressant recommendations that were for depression declined from 56% in 2005 to 51% in 2013. Mojtabai *et al.* (14) found that among physician visits where antidepressants were prescribed, the proportion of visits that did not have a psychiatric diagnosis increased from 60% in 1996 to 73% in 2007. Patten *et al.* (77) found that the number of physician recommendations for trazodone increased between 2000 and 2005 with an increasing percentage of these recommendations being for non-psychiatric reasons. For example, 14% of trazodone recommendations in 2000 were for sleep disorders compared to 20% in 2005. Finally, Meijer *et al.* (55) found that SSRIs that had been on the market longer were less likely to be prescribed for depression and more likely to be prescribed for other indications like anxiety disorders and social problems.

After depression, the most common treatment indications for antidepressants were anxiety disorders, sleeping disorders (e.g. insomnia), and pain. For anxiety disorders, five studies (17,31,32,74,76) found that the prevalence of this treatment indication was between 16% and

20%; however, one study (33) reported a prevalence of 27%, another study of nursing home residents (68) reported a prevalence of 6%, and another study of children and adolescents (65) also reported a prevalence of 6%. For sleeping disorders, three studies (17,32,68) reported a prevalence of 9% to 13% for this indication while another study (16) reported a slightly lower prevalence of 5%. For pain, three studies (17,31,66) only measured neuropathic pain and reported a similar prevalence of 1% to 2% for this indication. However, studies that used a broader definition of pain (e.g. chronic pain) reported a higher prevalence of this indication ranging from 6% (16) to 24% (32). Three studies reported the prevalence of multiple indications for antidepressant prescriptions, all with variable findings: 5% in a study using self-reported treatment indications from middle-aged to elderly antidepressant users (16), 12% in a study using physician-documented treatment indications for antidepressant users in a nursing home (68), and 40% in a study using diagnostic codes from administrative claims databases to measure treatment indications (31).

Finally, three studies estimated the percentage of antidepressant use for off-label indications. Eguale *et al.* (9) reported that 33% of antidepressant prescriptions were for off-label indications after comparing the physician-documented indications in the MOXXI system with the list of approved indications for the respective drug. In comparison, Larson *et al.* (30) reported that a lower percentage (19%) of antidepressant users were taking these drugs for off-label indications. However, their estimate may have been conservative because they defined off-label users as patients who had diagnostic codes for any condition within a pre-determined set of “off-label” indications, some of which were in fact on-label for certain antidepressants (e.g. diabetic neuropathy, fibromyalgia, insomnia, and premenstrual dysphoric disorder). On the other hand, Chen *et al.* (34) reported that 75% of antidepressant users were taking these drugs for off-label indications. However, their findings likely overestimated the true rate of off-label antidepressant use because patients without a diagnostic code for any approved indication were classified as off-label antidepressant users.

2.2 Limitations of previous studies

Studies that measured antidepressant treatment indications using diagnostic codes from administrative or electronic health record data often could not link prescriptions to an indication because the patient did not have diagnostic codes for any plausible conditions where antidepressants would be used. The percentage of antidepressant prescriptions with missing treatment indications ranged from 25% in one study using administrative claims data (32) to 59% in another study using electronic health record data (61). Moreover, we could not find any studies that assessed the accuracy of these diagnostic codes against a gold-standard measure of antidepressant treatment indications. Thus, although this method of measuring antidepressant treatment indications is simple and feasible to use with large databases, it is severely limited by the incompleteness of diagnostic codes and the lack of knowledge about their accuracy for determining antidepressant treatment indications.

In contrast, studies that used physician surveys or questionnaires to ascertain antidepressant treatment indications faced a different set of limitations. Many of these studies had poor response rates from physicians: 20% in one study (55), approximately 50% in two studies (52,71), and 68% (representing the median annual response rate over a 10 year period) in another study (14). Studies that measured treatment indications by interviewing physicians or patients included notably smaller numbers of antidepressant users (16,18,48,63) most likely because of the high cost and labour involved in conducting these interviews. Thus, although these methods for measuring antidepressant treatment indications may produce more valid results than relying on diagnostic codes, these methods are likely not feasible to use within the context of a large pharmacosurveillance system.

Indication-based electronic prescribing systems, on the other hand, are capable of accurately (37) and efficiently measuring treatment indications for large numbers of antidepressant users on an ongoing basis. When these systems are integrated into the routine process of care for patients, not only can they collect extremely valuable data for safety and effectiveness research (e.g. by identifying off-label uses and allowing us to learn from them), but they could also

benefit physicians by providing them with valuable feedback to improve medication safety (e.g. warnings about potential prescribing errors) and improve prescribing decisions (e.g. presenting physicians with the best therapeutic alternatives for a given indication) (11). However, such systems have not yet been widely adopted due to implementation challenges – namely designing a system that complements the current workflow of physicians and leverages information from other knowledge databases and technology systems (11). As a result, few studies have used data from such systems to perform pharmacovigilance activities for antidepressants. In this literature review, we identified only four studies to date that used data from medical record systems that linked medications directly to their physician-documented indications. Two of these studies involved medical record systems that were used more than 30 years ago by a very small number of physicians (49,78). The third study did not appear to require physicians to record treatment indications in the system as nearly one-third of the antidepressant prescriptions were missing treatment indications (56). And finally, although the fourth study (37) had complete data on physician-documented treatment indications for a large number of antidepressant prescriptions, the analysis did not focus specifically on antidepressants and thus lacked detail about the use of individual antidepressants for different clinical conditions and off-label indications.

2.3 Conclusions from the literature review

Most of the 41 studies identified in this literature review measured treatment indications for antidepressants using simple methods based on diagnostic codes that have not yet been validated or resource-intensive methods that are not feasible to use with large-scale pharmacosurveillance systems. Because of the different measurement methods used, it is not surprising that these studies reported a range of estimates for the prevalence of antidepressant use for different clinical conditions other than depression and off-label indications.

As an alternative to these methods, the use of indication-based electronic prescribing systems to measure treatment indications for antidepressant prescriptions represents an efficient and sustainable method that can be applied to large cohorts of antidepressant users. However,

because the widespread implementation of such systems is still a working area of research (11), we found that very few studies to date have used this method. Thus, until indication-based prescribing systems become more prevalent, it is a worthwhile endeavor to create predictive algorithms for antidepressant treatment indications that would allow researchers to accurately predict this information in the absence of their explicit documentation in administrative and electronic health record databases.

Table 2-1. Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Kreula and Hemminki, 1978 (47)	Reimbursements paid by the Social Insurance Office of Tampere for children <10 years of age in outpatient care in Finland	1974	All antidepressants under section 11d of the *Remedia Fennica*	Information available in the prescription	Among 42 antidepressant prescriptions, 17% were for enuresis while 76% had no indication recorded on the prescription.
Bergman *et al.*, 1979 (49)	Problem-oriented medical record system used at a multidoctor district health centre in northern Sweden (12 physicians)	1973	Tricyclic antidepressants	ICD-8 codes linked to prescribed drugs by secretaries according to the doctor's instruction, which were then transferred to the computer-based system	Among 257 antidepressant prescriptions, 51.8% were for depression mentis, 32.3% were for psychoneurosis, 5.8% were for manic-depressive psychosis, and 10% were for other diagnoses.
Bridges-Webb, Mant, and Hall, 1984 (78)	All patients of one practice in the Sydney University General Practice (3 physicians)	1978-1981	Amitriptyline	Computer-based monitoring system, where drugs prescribed were recorded and linked with the appropriate problem	Among 59 prescriptions for amitriptyline written in 1981, 47 (80%) were for neurotic depression.
Mackay *et al.*, 1997 (52)	Prescription Pricing Authority in England	The 1-year period following the release of each drug on the market (ranging from 1987 to 1992)	Four SSRIs: fluvoxamine, fluoxetine, sertraline, and paroxetine	Questionnaires sent to prescribing doctors at least 6 months following the prescription date	The response rate for questionnaires was about 50% for each drug. Among over 10,000 patients per drug, the percentage of patients prescribed the drug for depression ranged from 71.2% to 83.3%, for anxiety ranged from 1.5% to 7.9%, and for other indications ranged from 3.2% to 7.0%. No indication was specified for 11.1% to 15.5% of patients for each drug.

Abbreviations: SSRI = selective serotonin reuptake inhibitor

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Meijer *et al.*, 2001 (55)	109 psychiatrists (of 554 approached) in Utrecht, Netherlands	1995-1997	Four SSRIs: fluoxetine, fluvoxamine, paroxetine, and sertraline	Information obtained from the prescribing psychiatrist	Based on prescriptions for 1251 patients, SSRIs that were on the market longer were less likely to be prescribed for depression. Fluvoxamine was most often prescribed for anxiety disorder (36.8%), followed by fluoxetine (19.1%), paroxetine (16.9%), and sertraline (11.2%).
Ruths, Straand, and Nygaard, 2001 (57)	1552 residents at 23 nursing homes in Bergen, Norway	1997	All drugs under the ATC group N06A ("antidepressants")	Indications for drugs recorded by the nursing home physicians	Among 517 nursing home residents taking antidepressants, the 3 most common indications were depression (77%), restlessness (7%), and anxiety (5%).
Loosbrock *et al.*, 2002 (70)	US National Disease and Therapeutic Index physician survey (180 psychiatrists and 813 primary care physicians)	1997-1999	Citalopram, fluoxetine, paroxetine, sertraline, and venlafaxine	Indications recorded by physicians in confidential logbooks	Among 3 206 antidepressant prescriptions, 78.3% of prescriptions were for depression, 7.5% for anxiety, 2.2% for obsessive-compulsive disorder, 0.6% for eating disorders, 1.9% for stress, 0.9% for phobia, 0.9% for schizophrenia, 0.2% for obesity, and 7.6% for other indications.
Buhl Sørensen *et al.*, 2003 (58)	Cross-sectional survey of all Danish children and adolescents under 19 years of age taking an antidepressant	2000	SSRIs (citalopram, sertraline, paroxetine, fluoxetine, and fluvoxamine), TCAs (amitriptyline and nortriptyline), venlafaxine, mirtazapine, nefazodone, and moclobemide	Anonymous questionnaires set to all day-hospitals, acute in-patient units, out-patient units, and psychiatric out-patient child and adolescent clinics in Denmark	Among 382 patients, 49.6% of children were taking antidepressants for depression, 29.7% for obsessive-compulsive disorder, 10.6% for anxiety disorder, 6.2% for eating disorder, 0.8% for Tourette's syndrome, and 0.8% for hyperkinetic syndrome.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Henriksson *et al.*, 2003 (50)	Database containing purchases of prescription drugs for a representative sample of individuals in the county of Jämtland, Sweden	1995	TCAs and SSRIs	Access to patients' medical records, if the prescribing physician gave consent	Among 191 individuals with newly prescribed antidepressants, the medical record was consulted for 90% of these patients. Indications were found in the medical record for 95% of patients, with the most common indication being depression (66%), followed by anxiety and pain. By class, only 23% of TCAs were prescribed for depression, compared to 82% for SSRIs. TCAs were more commonly prescribed for anxiety and pain.
Pomerantz *et al.*, 2004 (71)	Patients enrolled in a single health maintenance organization in western Massachusetts who filled at least 1 new prescription for an antidepressant during the first 4 months of 2001	2001	SSRIs, TCAs, bupropion, trazodone, mirtazapine, venlafaxine, nefazodone	Surveys sent to prescribing physicians (51% response rate)	Among 485 patients with completed survey forms, 52% were prescribed antidepressants for depression or depression plus anxiety (39.2% for depression, 13.2% for depression plus anxiety), 11.8% for smoking cessation, 4.9% for migraine/headaches, 4.7% for chronic pain, 4.3% for fibromyalgia, 2.5% for premenstrual syndrome, 2.3% for insomnia, and 1.6% for hot flashes.
Chen *et al.*, 2006 (34)	Georgia Medicaid enrollees 18+ years old who filled at least 1 antidepressant prescription in 2001	2001	SSRIs, SNRIs, TCAs, bupropion, trazodone, mirtazapine, maprotiline, and nefazodone	ICD-9 codes recorded in claims data between January 2000 and December 2001	Among 62,289 patients who filled an antidepressant, 75% did not have a diagnostic code for an approved indication and were therefore classified as 'off-label recipients.'

Abbreviations: TCA = tricyclic antidepressant, SSRI = selective serotonin reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Petty *et al.*, 2006 (62)	Sample of patients from the North Bradford Primary Care Trust, UK (11 general practices)	2002-2004	Drugs under section 4.3 (antidepressants) of the British National Formulary	Diagnoses in general practitioner records and hospital letters	Among 140 patients prescribed an antidepressant, 44% were for depression, 17% for depression/anxiety, 7% for anxiety, 10% for pain, 4% for insomnia, and 9% for other indications, and 9% with no documented indication.
Larson, Miller, and Fleming, 2007 (30)	US MarketScan data	2002	SSRIs, TCAs, MAOIs, and bupropion	ICD-9 codes recorded in ambulatory and inpatient claims	Among 609 734 antidepressant users, 38.6% were using antidepressants for mental health or substance abuse disorder, 19.4% for off-label conditions, and 42.0% for unknown reasons.
Trifirò *et al.*, 2007 (66)	Arianna database of patients and general practitioners in Southern Italy (119 GPs)	2003-2004	All drugs under the ATC group N06A ("antidepressants")	Diagnoses recorded by physicians in the software dedicated to the Arianna Database	Among 11 418 patients prescribed an antidepressant, depression was the indication for 50.1% of TCA prescriptions, 67.2% of SSRI prescriptions, and 64.0% of prescriptions for other antidepressants. Anxiety was the indication for 13.8% of TCAs, 10.6% of SSRIs, and 10.7% of prescriptions for other antidepressants. No indication was recorded for 9.3% of TCA prescriptions, 3.7% of SSRI prescriptions, and 4.6% of prescriptions for other antidepressants.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant, MAOIs = monoamine oxidase inhibitors, ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Gardarsdottir *et al.*, 2007 (17)	Second Dutch National Survey of General Practice (195 GPs)	2001	SSRIs, TCAs, mianserin, mirtazapine, moclobemide, nefazodone, trazodone, tranylcypromine, and venlafaxine	International Classification of Primary Care (ICPC) codes recorded by GPs in the physician-patient contact file within 180-days around the index prescription date	Among 13 835 patients prescribed antidepressants, 45.5% were prescribed antidepressants for depression, 17.2% for anxiety/panic disorders, 9.3% for sleeping disorders, 1.7% for headache/migraine, 1.4% for obsessive-compulsive disorder, 1.2% for neuropathic pain, 0.3% for eating disorders, and 35.8% for unknown indication.
Patten, Esposito, and Carter, 2007 (77)	Canadian Disease and Therapeutic Index (652 physicians)	2000-2005	SSRIs, TCAs, venlafaxine, mirtazapine, bupropion, moclobemide, and trazodone	Anonymized diaries of treatment recommendations and reasons for prescribing recorded by physicians on two consecutive workdays per quarter	Between 2000 and 2005, prescribing of venlafaxine for depression and anxiety disorders increased; prescribing of SSRIs stayed relatively stable and most prescriptions were for depression; prescribing of trazodone increased with the proportion of prescriptions for depression declining (56.7% to 50.5%) but the proportion of prescriptions for sleep disorders increasing (14.0% to 20.0%); and prescribing of TCAs increased with the proportion of prescriptions for depression declining (32.6% to 19.4%). Overall, approximately 65% of all antidepressants were prescribed for depressive disorders.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Volkers, Heerdink, and Dijk, 2007 (56)	Children under 18 years of age in the Netherlands Information Network of General practice (97 practices in 2001, 73 practices in 2005)	2001 and 2005	All drugs under the ATC group N06A ("antidepressants")	Physician-documented indications recorded in the patient's medical record	2001: Among 732 antidepressant prescriptions for 194 patients, indications were recorded for 57.7% of prescriptions. Among patients 0-11 years, TCAs were mostly prescribed for nocturnal enuresis (84.4%) and 15.6% were for off-label indications. SSRIs were mostly prescribed for anxiety (53.8%) and 30.8% were for off-label indications. Among patients aged 12-17 years, TCAs were mostly prescribed for 'other' indications (52.3%) and 79.6% were for off-label indications. SSRIs were mostly prescribed for depression (70.6%) and 16.0% were for off-label indications. 2005: Among 458 prescriptions for 124 patients, indications were recorded for 60.9% of antidepressant prescriptions. Among patients 0-11 years, TCAs were mostly prescribed for nocturnal enuresis (93.4%) and 6.6% were for off-label indications. SSRIs were mostly prescribed for depression (54.5%) and 45.5% were for off-label indications. Among patients 12-17 years, TCAs were mostly prescribed for 'other' indications (38.7%) and other psychological problems (35.5%), and 74.2% were for off-label indications. SSRIs were mostly prescribed for depression (53.2%) and anxiety (22.8%) and 32.9% were for off-label indications.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant, ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Cascade, Kalali, and Thase, 2007 (73)	Verispan Prescription Drug and Diagnosis Audit database (3 100 office-based physicians from 29 specialties across the US)	2006-2007	SSRIs, SNRIs, TCAs, and 'other' antidepressants (e.g. bupropion and tetracyclics)	Physician surveys of indications for all drug mentions during one typical workday per month	SSRIs and newer antidepressants like venlafaxine and bupropion were primarily used to treat depression, anxiety, and bipolar disorders. 17% of duloxetine prescriptions were for pain compared to only 1% of prescriptions for other SSRIs and SNRIs. For amitriptyline, 46% of prescriptions were for neuropathic and muscular pain conditions, 8% were for migraine, while only 25% were for CNS conditions.
Sihvo *et al.*, 2008 (18)	Cross-sectional population-based Finnish Health 2000 Study	2000-2002	All drugs under the ATC group N06A ("antidepressants")	Structured interviews with patients and information on past psychiatric hospitalizations from the National Hospital Discharge Register	Among 526 antidepressant users, 59% had depression or a history of depression, and 27% had anxiety disorder or a history of anxiety. Approximately 25% of users had no known psychiatric morbidity.
Chermá *et al.*, 2008 (51)	Elderly residents in 8 nursing homes in the county of Östergötland, Sweden	2003	All drugs under the ATC group N06A ("antidepressants")	Medical records forms at the nursing homes	Among 71 individuals on antidepressants, indications were identified for 96% of individuals. Depression was the indication for 60% of patients and dystimi for 16% of patients.
Zullino *et al.*, 2008 (67)	Psychiatric hospital of the University of Lausanne, Switzerland	6 reference days from April 1999 to November 2001	SSRIs, TCAs, and mirtazapine, reboxetine, nefazodone, and moclobemide	ICD-10 codes recorded for patients on the reference day	Among 174 patients prescribed antidepressants, 46.6% were prescribed antidepressants for depression, 5.7% for anxiety disorder, and 47.7% for other indications.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor, TCA = tricyclic antidepressant, ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Gardarsdottir *et al.*, 2009 (59)	Second Dutch National Survey of General Practice (derivation set); Integrated Primary Care Information database (validation set)	2001	SSRIs, TCAs, mianserin, mirtazapine, moclobemide, nefazodone, oxitriptan, phenelzine, trazodone, tranylcypromine, and venlafaxine	International Classification of Primary Care (ICPC) codes recorded by GPs in the patient's medical file in the 12 months around the index prescription date	Among 1 855 new antidepressant users in the derivation set, 51.8% had a diagnostic code for depression. Among 3 231 new antidepressant users in the validation set, 46.2% had a diagnostic code for depression.
Haw and Stubbs, 2010 (63)	Adolescents under 18 years of age in two medium secure inpatient units at St. Andrew's Hospital in Northampton, UK	2007	Not specified, but mostly SSRIs	Interviews with patients' consultant psychiatrist	Among 89 patients, 25 patients were taking antidepressants, 44% of which were taking the antidepressant off-label
Milea *et al.*, 2010 (31)	PharMetrics US administrative claims database	2003-2004	Bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, and venlafaxine	ICD-9 code for an approved or clinically-accepted diagnosis for antidepressant treatments within the month before or after the index claim date	Among 392 409 new antidepressant users, 53.3% did not have a diagnosis of interest recorded. Among patients with a diagnosis of interest, 29.5% had a diagnostic code for depression, 17.4% for anxiety disorders, 4.5% for abuse and dependence, 2.5% for disorders specific to childhood and adolescence (e.g. ADHD), 1.9% for fibromyalgia, 1.3% for pain, 1.3% for bipolar disorders, 0.9% for premenstrual disorders, 0.6% for obsessive compulsive disorders, and 0.5% for eating disorders. 74.6% of patients with a diagnosis of interest had one condition recorded, 21.9% had two, and 3.5% had three or more.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Mark, 2010 (74)	US National Disease and Therapeutic Index (approximately 4000 physicians)	2005	TCAs, tetracyclics, MAOIs, SSRIs, SNRIs, and newer generations antidepressants	Indications recorded by physicians in confidential logbooks	Among 66 855 drug mentions, 92.7% were for psychiatric conditions, which included mood disorders like depression (65.3%), anxiety disorders (16.4%), attention-deficit disorders (2.8%), schizophrenia and other psychotic disorders (2.6%), and adjustment disorders (1.3%). Non-psychiatric conditions included headache/migraine (1.1%), other connective tissue disease (1.0%), other nervous system disorders (0.8%), other female genital disorders (0.8%), and spondylosis/other back problems (0.7%).
Mojtabai and Olfson, 2011 (14)	US National Ambulatory Medical Care Surveys	1996-2007	Not specified	Diagnoses recorded for the index visit where the antidepressant was prescribed	Annual survey response rate ranged from 62.9% to 77.1% during the study period. Psychiatric diagnoses were responsible for prescriptions in only 44.0% of antidepressant visits with primary care providers and only 12.8% of antidepressant visits with other nonpsychiatrist providers. The proportion of antidepressant visits lacking a psychiatric diagnosis increased from 59.5% in 1996 to 72.7% in 2007.

Abbreviations: TCA = tricyclic antidepressant, MAOI = monoamine oxidase inhibitors, SSRI = selective serotonin reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Harris *et al.*, 2012 (53)	Patients aged 65 years or older in The Health Improvement Network (326 general practices in England and Wales)	2008-2009	TCAs, MAOIs, SSRIs, SNRIs, mirtazapine, flupentixol, mianserin, reboxetine, trazodone, and tryptophan	Read codes for depression, anxiety, and other possible indications in patients' medical records	Among 28 762 older community-dwelling patients, 63.6% had a diagnosis of depression ever recorded, 9.1% had a diagnosis of depression in the last year, and 28.3% had no history of depression or diagnosis of depression or anxiety symptoms in the last year. Among 2 955 older patients in care homes, 47.9% had a diagnosis of depression ever recorded, 4.0% had a diagnosis of depression in the last year, and 42.5% had no history of depression or diagnosis of depression or anxiety symptoms in the last year.
Lee *et al.*, 2012 (75)	Cross-sectional analysis of children and adolescents aged 6-18 years from the US National Ambulatory Medical Care Survey	2000-2006	32 US FDA-approved antidepressants including SSRIs, TCAs, MAOIs, SNRIs, and serotonin antagonist reuptake inhibitors	Diagnoses recorded for the index visit where the antidepressant was prescribed	Among 1 170 office visits where antidepressants were prescribed, only 9.2% of antidepressants were associated with an FDA-approved indication for the patient, which included the indications depression, obsessive-compulsive disorder, and nocturnal enuresis.
Wu *et al.*, 2012 (32)	National Health Insurance Research Database in Taiwan	2000-2009	All drugs under the ATC group N06A ("antidepressants")	ICD-9 codes recorded on or within 90 days after the first prescription date	Among 156 125 incident antidepressant users, 21.3% had diagnostic codes for mood disorders, 18.1% for anxiety disorders, 12.9% for sleep disorders, and 23.9% for chronic pain. Treatment indications could not be identified for 25.4% of antidepressant users.

Abbreviations: TCA = tricyclic antidepressant, MAOI = monoamine oxidase inhibitors, SSRI = selective serotonin reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor, FDA = US Food and Drug Administration, ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Eguale *et al.*, 2012 (9)	The Medical Office of the XXI Century electronic health record network in Quebec, Canada (113 primary care physicians)	2005-2009	Drugs classified as antidepressants under the American Hospital Formulary Service	Indications documented by the prescribing physician on the electronic prescription	Among approximately 16 000 new prescriptions for antidepressants, 33.4% were for an off-label indication.
Bourgeois *et al.*, 2012 (68)	The Prescribing in Homes for the Elderly in Belgium (PHEBE) study	2006	Drugs under the following ATC classes: N06AA, N06AB, N06AG, and N06AX	Questionnaires sent to residents' GPs with a checklist of possible indications for the antidepressant	Among 551 antidepressant users taking a single antidepressant, 66.2% were taking the antidepressant for depression, 13.5% for insomnia, 6.2% for anxiety, and 1.6% for pain. 11.6% of residents were taking a single antidepressant for multiple indications.
Steinhausen *et al.*, 2014 (60)	Children and adolescents under 18 years of age in the Danish prescription register	1996-2010	All drugs under the ATC group N06A ("antidepressants")	ICD-10 codes for patients recorded in the same year in the Danish Psychiatric Central Register (DPCR)	Increase in dispensed prescriptions for antidepressants over time was matched by an increase in the prevalence of depression, obsessive-compulsive disorders, and adjustment disorders among patients. Analysis of indications had to be limited to 3 major indications for antidepressants because the prescription register did not contain data on diagnoses and diagnoses had to be obtained from the DPCR instead.

Abbreviations: ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Simon *et al.*, 2014 (33)	Records from 10 large US integrated health systems in the Mental Health Research Network	2010	All drugs approved by the FDA for major depressive disorder plus several drugs with similar chemical and clinical effects (e.g. fluvoxamine and clomipramine)	Outpatient and inpatient ICD-9 codes for possible mental health diagnoses for antidepressants recorded in electronic medical record and claims data in 2010	Among 1 011 946 health plan members taking antidepressants, 48% had a diagnostic code for depression, 27% for anxiety disorder, 3% for bipolar disorder, and 3% for attention deficit disorder. 39% had no mental health diagnosis recorded.
Abbing-Karahagopian *et al.*, 2014 (69)	Seven electronic healthcare databases from five European Countries	2001-2009	SSRIs and TCAs	Diagnostic codes (ICPC, Read, or ICD-9 codes) recorded within a) 3 months before and after the first prescription, and b) any time before the first prescription, up to 1 January 2001	Using a window of 3 months around the index prescription date, between 12.1% and 57.1% of prescriptions were associated with a diagnosis for depression. This range increased to 39.0% to 69.0% when diagnostic codes recorded back to 1 January 2001 were considered.
Mercier *et al.,* 2014 (48)	50 patients belonging to 28 general practitioners from a local research network in the Normandy region of France	2011-2012	SSRIs and SNRIs	Face-to-face interviews with prescribing physicians	Among the 50 patients assessed, 9 had no psychiatric diagnosis.

Abbreviations: FDA = Food and Drug Administration, SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant, SNRI = serotonin-norepinephrine reuptake inhibitor, ICPC = International Classification of Primary Care

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Noordam *et al.*, 2015 (61)	Patients aged 10 years or older in the Dutch Integrated Primary Care Information database	1996-2012	All drugs under the ATC group N06A ("antidepressants")	International Classification of Primary Care (ICPC) codes for selected indications recorded for the patient within 90 days before or after the first antidepressant prescription	Diagnoses of interest were recorded for 41% of incident prescriptions, decreasing from 68% in 1996-1997 to 40% in 2012. Over the study period, antidepressants were increasingly prescribed for sleep disorders and neuropathic pain.
Hassan *et al.*, 2016 (54)	Cross-sectional study of prisoners from 11 prisons in England	2012-2013	Drugs under section 4.3 (antidepressants) of the British National Formulary	Information in patients' clinical notes	Among men, 61.8% of antidepressants were prescribed for a valid indication. Among women, 76.8% of antidepressants were prescribed for a valid indication.
Aarts *et al.*, 2016 (16)	Individuals older than 45 years of age in the Rotterdam Study cohort	1997-2013	All drugs under the ATC group N06A ("antidepressants")	Patient interviews (self-report)	Among 914 antidepressant users, 52.4% reported taking antidepressants for depression, 8.8% for anxiety, 13.0% for stress, 4.9% for sleep disorders, 1.1% for headache/migraine, 5.9% for pain, and 13.5% for 'unknown' indications. Approximately 5% of users reported taking the antidepressant for multiple indications.
John *et al.*, 2016 (64)	Individuals aged 6-18 years in the Secure Anonymized Information Linkage System (SAIL Databank) in Wales, UK	2003-2013	Drugs under section 4.3 (antidepressants) of the British National Formulary	Read codes recorded in the General Practice Database in the year before and 6 months after the incident prescription	Over 50% of new prescriptions were associated with a diagnostic code for depression. Plausible indications could not be found for 16.8% of new prescriptions.

Abbreviations: ATC = Anatomical therapeutic chemical classification system

Table 2-1. (continued) Previous studies (n=41) that have measured treatment indications for antidepressants

Authors	Population/ Data source	Data collection period	Antidepressants included	Methodology for measuring treatment indications	Findings
Sarginson *et al.*, 2017 (65)	Children aged 3 to 17 in the UK Clinical Practice Research Datalink (CPRD)	2000-2015	SSRIs, TCAs, and 'other' antidepressants	Diagnostic codes for depression or possible alternative uses recorded in the CPRD between 12 weeks prior and 4 weeks after the incident prescription	In 2015, 21.4% of new prescriptions were associated with a depression diagnosis, 6.0% for anxiety, 4.3% for pain, 2.3% for headache, 1.4% for obsessive-compulsive disorder, 0.9% for enuresis, and 0.8% for eating disorders. 43.6% of patients could not be assigned a treatment indication.
O'Brien *et al.*, 2017 (76)	US National Disease and Therapeutic Index (4 120 physicians in 2005; 4 140 physicians in 2013)	2005 and 2013	Antidepressants under the 2014 Uniform System of Classification	Indications recorded by physicians in confidential logbooks	The proportion of patients with depression that were prescribed antidepressants declined from 56.4% in 2005 to 50.8% in 2013. However, the proportion of patients with anxiety disorders that were prescribed antidepressants increased from 16.4% in 2005 to 19.8% in 2013 and also increased over time for attention-deficit disorders, connective tissue diseases, headache/migraine, and osteoarthritis.

Abbreviations: SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressant

3 Data source

3.1 The Medical Office of the XXIst Century (MOXXI)

The Medical Office of the XXIst Century (MOXXI) is an electronic prescribing and drug management system equipped with several functionalities to improve the safety and quality of care for patients, particularly around medication use. Among these functionalities, the MOXXI system features an electronic prescribing tool that requires physicians to document treatment indications for all prescriptions written. The presence of prescriptions explicitly linked to their corresponding medical problems is the key feature that makes the MOXXI system an ideal data source for addressing the research objectives in this dissertation.

3.1.1 Recruitment

Developed by Dr. Robyn Tamblyn and colleagues from McGill University, the third and most recent phase of the MOXXI system (MOXXI III) was rolled out in 2003. Annual licensure-renewal data from the Collège des médecins du Québec were used to identify eligible physicians to participate in the MOXXI III research program. Physicians were eligible if they were full-time primary care physicians working in office-based, fee-for-service practices located in selected metropolitan areas in Montreal and Québec City (36). Patients were eligible to participate if they had visited a MOXXI physician in the past year. Eligible patients were identified from medical billing, physician, and beneficiary files from the RAMQ and were added to an electronic patient list that was provided to each physician to prepopulate their practice population in the MOXXI system (36). Recruitment of MOXXI physicians and patients has been ongoing since 2003 with approximately 185 physicians (25% of eligible) and 100 000 patients (30% of eligible) currently participating in the research program. All MOXXI physicians and patients have given consent to have their information used for research purposes.

3.1.2 Characteristics of MOXXI physicians and patients

Compared to non-consenting physicians, MOXXI physicians are on average five years younger. Among MOXXI physicians, physicians are more likely to recruit patients and use the system if

they are more technologically proficient, have lighter patient workloads, and see patients with less fragmented care (80). Patients are more likely to consent to participate in MOXXI if they are older, have more complex health issues, have higher income, and visit their study physician more often (80).

3.2 Quebec's administrative health databases

Many of the unique functionalities of the MOXXI system are possible because of the system's integration with databases from Quebec's health insurance agency, the Régie de l'assurance maladie du Québec (RAMQ), and the provincial hospital discharge summary database, MED-ECHO. Linked health services information from these data sources were used in this dissertation, particularly in objectives 2-4.

3.2.1 The Régie de l'assurance maladie du Québec (RAMQ)

The RAMQ manages the health insurance of all Quebec residents (99% of the Quebec population) and maintains a database with the name, age, sex, and residence of all beneficiaries. The RAMQ also manages the reimbursement of all physicians and pharmacies. For approximately 80% of physicians in the province who work under a fee-for-service payment structure (81), a reimbursement claim is submitted to RAMQ for each medical service performed. Each claim captures information about the date and location of the visit, type of service performed, and an optional diagnostic code representing the main reason for the visit. The RAMQ also maintains a database of all medications dispensed by pharmacies in Quebec to beneficiaries enrolled in the public drug insurance program. Approximately 50% of Quebec residents are registered in the public drug insurance program including the elderly, welfare recipients, and persons not insured through their employer (36).

3.2.2 MED-ECHO

The MED-ECHO database, maintained by the Quebec Ministry of Health and Social Affairs, collects information on all hospitalizations occurring at acute care institutions in the province (82). Information about the patient's age, sex, admission and discharge dates, primary and

secondary diagnoses, and medical procedures performed in hospital are recorded in the database based on structured chart abstraction conducted by medical archivists.

3.3 Antidepressants and their licensed indications

In this dissertation, antidepressants were defined as drugs that were approved for depression by Health Canada. Table 3-1 lists the 25 drugs that were classified as antidepressants along with their date of market approval in Canada and their licensed indications by either Heath Canada or the US Food and Drug Administration (FDA) as of September 30, 2015. Approved indications for antidepressants in both Canada and the USA were identified to yield more conservative estimates of off-label prescribing and because the prescribing practices of physicians in Quebec may be influenced by pharmaceutical advertising from the USA. The market approval dates for antidepressants were obtained through email correspondence with a representative from Health Canada. Licensed indications for antidepressants in Canada and the USA were obtained from drug monographs in the Vigilance knowledge database (83) and the US DrugDex compendium (84), respectively.

Four antidepressants (escitalopram, duloxetine, desvenlafaxine, and vortioxetine) obtained market approval after 2003 – the year when the MOXXI III system was implemented. Another antidepressant, nefazodone, was discontinued in Canada in November 2003 (Table 3-1). SSRIs and SNRIs had notably more licensed indications than other types of antidepressants. Within the class of SSRIs and SNRIs, older drugs generally had more licensed indications than newer drugs. The list of licensed indications for each drug according to Health Canada and the FDA was often the same. However, there were slight discrepancies for 10 drugs. Five of these drugs (escitalopram, fluvoxamine, clomipramine, nortriptyline, and moclobemide) had indications licensed by Health Canada only, while another five drugs (fluoxetine, sertraline, doxepin, imipramine, and maprotiline) had indications licensed by the FDA only (Table 3-1).

3.4 Cohort creation

3.4.1 Accrual period for antidepressant prescriptions

The start of the accrual period for antidepressant prescriptions was January 1, 2003 in all manuscripts except manuscript 1 (Table 3-2). Because manuscript 1 measured trends over time in the proportion of antidepressants prescribed for depression, prescriptions written before 2006 were not included in the analysis because there were far fewer antidepressant prescriptions written during these years (Table 3-3). The lower numbers of antidepressant prescriptions written before 2006 was because the MOXXI system was being used by fewer physicians, all of whom practiced in Montreal. Around 2006, however, more physicians were recruited to participate in MOXXI and the study population expanded to include patients in Québec City.

The accrual period for antidepressant prescriptions ended on September 30, 2015 in the first two manuscripts and December 31, 2012 in the remaining three manuscripts (Table 3-2). The later end date in the first two manuscripts was because these studies did not require information from the RAMQ or MED-ECHO databases in their analyses. Thus, prior to the publication of manuscript 1 in 2016, I updated the analysis to include additional prescriptions. For the remaining manuscripts, however, the analysis used variables derived from complex data tables that were assembled in 2013/2014 based on static data files from RAMQ and MED-ECHO that contained information only up to December 31, 2012.

3.4.2 Exclusion criteria

The manuscripts in this dissertation used slightly different exclusion criteria to match the objectives of their corresponding analysis (Table 3-2). Manuscripts 1 and 2, which investigated off-label indications for antidepressant prescriptions, excluded prescriptions for patients under 18 years of age because many of the antidepressants and their licensed indications in Table 3-1 are not approved for children and adolescents. Moreover, there were not enough antidepressant prescriptions in the MOXXI system to perform a robust analysis on off-label

indications for children and adolescents. In fact, there were only 229 (0.3%) antidepressant prescriptions for patients under 18 years of age between January 2003 and September 2015.

Prescriptions for monoamine oxidase inhibitors (MAOIs) were excluded in manuscript 1 because the paper reported a stratified analysis of risk differences over time by drug class and there were too few prescriptions for MOAIs (n=133) to perform a reliable analysis (Table 3-2). Likewise, antidepressants that were prescribed on average less than once per month during the study period were excluded in three other manuscripts either because the study reported results by individual drug (manuscript 2) or to avoid running multivariable models with sparse cells (manuscripts 4 and 5).

Because manuscripts 3 to 5 required information from the RAMQ and MED-ECHO databases, these manuscripts excluded patients whose administrative health data could not be obtained due to faulty or failed linkages. These three manuscripts also excluded antidepressant prescriptions that were written at least one week outside the period during which the patient consented to participate in MOXXI. These prescriptions were excluded because all records for the patient in RAMQ and MED-ECHO databases over the past year may not have been pulled.

Finally, in manuscripts 4 and 5, prescriptions were excluded if they had missing values for any of the study covariates. Missing data occurred only if the patient's postal code was outside a census tract area or if the prescriber had not completed the evidence-nonconformity-practicality survey (85) at the time he or she was recruited to participate in MOXXI. Thus, because only a small (5.0%) proportion of prescriptions had missing data, these prescriptions were excluded from the analysis rather than employing multiple imputation. Moreover, it is highly unlikely that the mechanisms behind the missing data are related to factors affecting the medical reasons why antidepressants are prescribed.

Table 3-2 shows the final number of antidepressant prescriptions that were used in each manuscript after applying all exclusion criteria. In all manuscripts, the unit of the analysis was the prescription.

Table 3-1. Drugs approved for depression in Canada

Class	Drug (molecule name)	Year approved in Canada[a]	Licensed indications by Health Canada or the US Food and Drug Administration[b]	
			No. indications	Indications
SSRI	Fluoxetine	1988	5	Depression, obsessive-compulsive disorder, bulimia, panic disorders[c], premenstrual disorders[c]
	Fluvoxamine	1990	2	Depression[d], obsessive-compulsive disorder
	Sertraline	1992	6	Depression, obsessive-compulsive disorder, panic disorder, post-traumatic stress disorder[c], social phobia[c], premenstrual disorders[c]
	Paroxetine	1993	7	Depression, anxiety, obsessive-compulsive disorder, panic disorders, post-traumatic stress disorder, premenstrual disorders, social phobia
	Citalopram	1999	1	Depression
	Escitalopram	2004	3	Depression, anxiety, obsessive-compulsive disorder[d]
SNRI	Venlafaxine	1994	4	Depression, anxiety, panic disorders, social phobia
	Duloxetine	2007	5	Depression, anxiety, chronic pain, diabetic neuropathy, fibromyalgia
	Desvenlafaxine	2009	1	Depression

Abbreviations: SSRI = selective serotonin reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor

[a]Approval years for antidepressants were obtained via email correspondence with a representative from Health Canada

[b]Licensed indications by Health Canada were obtained from drug monographs in the Vigilance database. Licensed indications by the US Food and Drug Administration were obtained from drug monographs in the DrugDex compendium.

[c]Indications approved for the drug by the US Food and Drug Administration only

[d]Indications approved by Health Canada only

Table 3-1. (continued) Drugs approved for depression in Canada

Class	Drug (molecule name)	Year approved in Canada[a]	Licensed indications by Health Canada or the US Food and Drug Administration[b]	
			No. indications	Indications
TCA	Imipramine	1959	1	Depression, enuresis[c]
	Nortriptyline	1963	4	Depression, enuresis[d], nicotine addiction[d]
	Trimipraine	1963	1	Depression
	Desipramine	1965	1	Depression
	Amitriptyline	1966	1	Depression
	Doxepin	1970	4	Depression, anxiety, insomnia, pruritus[c]
	Clomipramine	1973	2	Depression[d], obsessive-compulsive disorder
MAOI	Phenelzine	1959	1	Depression
	Tranylcypromine	1960	2	Depression, depressive state in bipolar disorder
	Moclobemide	1992	1	Depression[d]
Other	Maprotiline	1976	3	Depression, bipolar disorder[c], mixed anxiety and depression[c]
	Trazodone	1982	1	Depression
	Nefazodone[e]	1994	1	Depression
	Bupropion	1998	2	Depression, nicotine addiction
	Mirtazapine	2001	1	Depression
	Vortioxetine	2014	1	Depression

Abbreviations: TCA = tricyclic antidepressant, MAOI = monoamine oxidase inhibitor

[a]Approval years for antidepressants were obtained via email correspondence with a representative from Health Canada

[b]Licensed indications by Health Canada were obtained from drug monographs in the Vigilance database. Licensed indications by the US Food and Drug Administration were obtained from drug monographs in the DrugDex compendium.

[c]Indications approved for the drug by the US Food and Drug Administration only

[d]Indications approved by Health Canada only

[e]Nefazodone was discontinued in Canada in November 2003

Table 3-2. Criteria used to create the dataset of antidepressant prescriptions in each manuscript

Manuscript	Accrual period for antidepressant prescriptions in the MOXXI system	Exclusion criteria (no. prescriptions removed)	Final no. prescriptions
1	January 1, 2006 to September 30, 2015 (n=102,153)	1. Patients <18 years old (n=261) 2. Prescriptions for monoamine oxidase inhibitors (n=133)	101,759
2	January 1, 2003 to September 30, 2015 (n=107,368)	1. Patients <18 years old (n=289) 2. Antidepressants with <150 prescriptions[a] (n=229)	106,850
3	January 1, 2003 to December 31, 2012 (n=79,134)	1. Patients who could not be linked to RAMQ (n=903) 2. Prescriptions written >1 week outside the consent period (n=531)	77,700
4 and 5	January 1, 2003 to December 31, 2012 (n=79,134)	1. Patients who could not linked to RAMQ (n=903) 2. Prescriptions written >1 week outside the consent period (n=531) 3. Antidepressants with <120 prescriptions[b] (n=147) 4. Prescriptions with missing covariate information (n=3977)	73,576

Abbreviations: MOXXI = The Medical Office of the XXIst Century

[a]Excluded prescriptions for six drugs: moclobemide (n=122), vortioxetine (n=56), maprotiline (n=18), phenelzine (n=13), tranylcypromine (n=10), and nefazodone (n=10)

[b]Excluded prescriptions for five drugs: moclobemide (n=96), maprotiline (n=18), phenelzine (n=13), tranylcypromine (n=10), and nefazodone (n=10)

Table 3-3. Number of antidepressant prescriptions in the MOXXI system, by calendar year and region

Calendar year	Number of antidepressant prescriptions identified in the MOXXI system (%[a])		
	Overall	By region	
		Montreal	Québec City
2003	1488	1488 (100.0)	0 (0.0)
2004	1462	1462 (100.0)	0 (0.0)
2005	2265	2000 (88.3)	265 (11.7)
2006	7586	2840 (37.4)	4746 (62.6)
2007	9250	3500 (37.8)	5750 (62.2)
2008	9931	4218 (42.5)	5713 (57.5)
2009	11350	4419 (39.6)	6859 (60.4)
2010	11994	4631 (38.6)	7363 (61.4)
2011	12155	4635 (38.1)	7520 (61.9)
2012	11653	4390 (37.7)	7263 (62.3)
2013	11367	4014 (35.3)	7353 (64.7)
2014	10019	3909 (39.0)	6110 (61.0)
2015[b]	6848	3302 (48.2)	3546 (51.8)
TOTAL	**107,368**	**44,880 (41.8)**	**62,488 (58.2)**

Abbreviations: MOXXI = The Medical Office of the XXIst Century

[a]Using the total number of antidepressant prescriptions written in any location during that calendar year as the denomiator

[b]Incudes prescriptions written to September 30, 2015 only

4 Treatment indications for antidepressants prescribed in primary care in Quebec, Canada, 2006-2015

4.1 Preamble

In this paper, I performed a detailed analysis of treatment indications for antidepressant prescriptions to motivate the need for heightened pharmacovigilance around antidepressants, thus setting the stage for the later manuscripts in this thesis.

In the analysis, I used nearly 10-years of electronic prescribing data from the MOXXI system to measure the relative frequency of different clinical indications for antidepressant prescriptions and trends in antidepressant prescribing for depression versus other indications over the past decade. I also measured the proportion of antidepressant prescription that were written for an indication that was unapproved or "off-label" for the prescribed drug.

These findings were published as **research letter** in *JAMA* in May 2016. Appendix C contains a reprint of the original article.

4.2 Title page and footnotes

Title: Treatment indications for antidepressants prescribed in primary care in Quebec, Canada, 2006-2015

Authors: Jenna Wong[1], MSc; Aude Motulsky[2], PhD; Tewodros Eguale[3], MD, PhD; David L Buckeridge[1], MD, PhD; Michal Abrahamowicz[1], PhD; and Robyn Tamblyn[1], PhD

Affilitations:

[1]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

[2] Department of Management, Evaluation and Health Policy, School of Public Health, Université de Montréal

[3]School of Pharmacy, Massachusetts College of Pharmacy and Health Sciences (MCPHS), Boston, MA

Corresponding author:

Jenna Wong

1140 Pine Avenue West

Montreal, Quebec, Canada

H3A 1A3

Email: jenna.wong@mail.mcgill.ca

Citation: 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.

4.3 Introduction

Antidepressant use in the United States has increased over the last 2 decades (86). A suspected reason for this trend is that primary care physicians are increasingly prescribing antidepressants for non-depressive indications, including unapproved ('off-label') indications that have not been evaluated by regulatory agencies (17). However, the frequency with which physicians prescribe antidepressants for non-depressive indications is unknown because treatment indications are rarely documented. We analyzed the prevalence of treatment indications for antidepressants and assessed temporal trends in antidepressant prescribing for depression.

4.4 Methods

This study used data from the MOXXI research platform(36). MOXXI is an electronic medical record (EMR) and prescribing system that has been used by primary care physicians in community-based, fee-for-service practices around 2 major urban centers in Quebec, Canada. During the study period, approximately 185 physicians (25% of eligible) and 100,000 patients (30% of all who visited a MOXXI physician) gave informed consent to use the EMR and have their information used for research purposes. Compared to non-consenters, MOXXI physicians were younger and MOXXI patients were older with more health complexities (80).

This study included all prescriptions written for adults between 1 January 2006 and 30 September 2015 for all antidepressants except monoamine oxidase inhibitors. Physicians had to document at least 1 treatment indication per prescription using a drop-down menu containing a list of indications or by typing the indication(s). In a validation study, these indications had excellent sensitivity (98.5%) and high positive predictive value (97.0%) (37). Prescriptions were classified as on-label or off-label depending on whether the drug was approved for the indication by Health Canada or the US Food and Drug Administration by September 2015. Temporal trends in antidepressant prescribing for depression were measured using generalized linear risk difference models for binary outcomes, with an identity link. A linear effect of calendar time (in years) was modeled on the probability of antidepressant prescribing for depression, adjusted for patient age and sex and accounting for multi-level clustering of

prescriptions using an alternating logistic regression algorithm (87). All statistical analyses were conducted using SAS software (version 9.4). This study was approved by the McGill Institutional Review Board.

4.5 Results

During the study period, 101,759 antidepressant prescriptions (5.9% of all prescriptions) were written by 158 physicians for 19,734 patients. Only 55.2% of antidepressant prescriptions were indicated for depression. Physicians also prescribed antidepressants for anxiety disorders (18.5%), insomnia (10.2%), pain (6.1%) and panic disorders (4.1%) (Table 4-1). For these indications, respectively, the most frequently prescribed antidepressants were citalopram (29.5% of prescriptions for the indication), trazodone (76.6%), amitriptyline (65.1%), and paroxetine (35.9%).

For 29.4% of all antidepressant prescriptions (65.6% of prescriptions not for depression), physicians prescribed a drug for an off-label indication, especially insomnia and pain. Physicians also prescribed antidepressants for several indications that were off-label for all antidepressants, including migraine, vasomotor symptoms of menopause, attention deficit hyperactivity disorder, and digestive system disorders (Table 4-1).

Between 2006 and 2015, the percentage of antidepressants prescribed for depression decreased significantly, with an adjusted 5-year difference of -9.73% (95% CI -11.86% to -7.61%) for serotonin-norepinephrine reuptake inhibitors, -3.96% (-5.33% to -2.59%) for selective serotonin reuptake inhibitors, and -2.99% (-4.90% to -1.08%) for tricyclic antidepressants (Figure 4-1). However, the percentage of ‘other’ antidepressants (especially mirtazapine) prescribed for depression increased significantly (adjusted 5-year difference of 2.36%, 0.32% to 4.40%).

4.6 Discussion

Between 2006 and 2015, primary care physicians in Quebec commonly and increasingly prescribed antidepressants for non-depressive indications. When physicians prescribed antidepressants for insomnia and pain, they often prescribed antidepressants off-label.

The study was limited by a selective patient population and a small number of prescribers from 1 Canadian province. However, this is the first study to our knowledge to describe the prevalence of treatment indications for antidepressants using validated, physician-documented treatment indications recorded at the point of prescribing. The findings indicate that the mere presence of an antidepressant prescription is a poor proxy for depression treatment and they highlight the need to evaluate the evidence supporting off-label antidepressant use.

Table 4-1. Treatment indications and off-label prescribing for antidepressant prescriptions in Quebec, Canada, 2006-2015

Treatment indication[a]	No. of prescriptions (%)[b]	N (%) by pharmacological class[c]				N (%) off-label[h]
		SSRI[d]	SNRI[e]	TCA[f]	Other[g]	
Depressive disorders	56,154 (55.2)	26,339 (46.9)	15,259 (27.2)	1,502 (2.7)	13,054 (23.3)	0 (0.0)
Anxiety disorders[i]	18,849 (18.5)	12,466 (66.1)	5,076 (26.9)	273 (1.5)	1,034 (5.5)	8,975 (47.6)
Insomnia	10,347 (10.2)	19 (0.2)	2 (0.0)	2,242 (21.7)	8,804 (78.1)	10,077 (97.4)
Pain	6,241 (6.1)	24 (0.4)	1,340 (21.5)	4,623 (74.1)	254 (4.1)	5,174 (82.9)
Panic disorders with or without agoraphobia	4,174 (4.1)	3,280 (78.6)	751 (18.0)	89 (2.1)	54 (1.3)	1,487 (35.6)
Fibromyalgia	1,550 (1.5)	63 (4.1)	958 (61.8)	506 (32.7)	23 (1.5)	946 (61.0)
Migraine	1,498 (1.5)	6 (0.4)	22 (1.5)	1,470 (98.1)	0 (0.0)	1,498 (100.0)
Obsessive-compulsive disorder	1,111 (1.1)	875 (78.8)	177 (15.9)	53 (4.8)	6 (0.5)	435 (39.2)
Vasomotor symptoms of menopause	856 (0.8)	112 (13.1)	736 (86.0)	2 (0.2)	6 (0.7)	856 (100.0)
Social phobia	568 (0.6)	434 (76.4)	134 (23.6)	0 (0.0)	0 (0.0)	199 (35.0)
Nicotine dependence	514 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	514 (100.0)	0 (0.0)
Attention deficit hyperactivity disorder	389 (0.4)	16 (4.1)	4 (1.0)	5 (1.3)	364 (93.6)	389 (100.0)
Post-traumatic stress disorder	263 (0.3)	211 (80.2)	35 (13.3)	2 (0.8)	15 (5.7)	207 (78.7)
Sexual dysfunction	261 (0.3)	39 (14.9)	2 (0.8)	8 (3.1)	212 (81.2)	261 (100.0)
Premenstrual disorders and syndromes	212 (0.2)	193 (91.0)	17 (8.0)	2 (0.9)	0 (0.0)	63 (29.7)
Digestive system disorders	119 (0.1)	4 (3.4)	1 (0.8)	114 (95.8)	0 (0.0)	119 (100.0)
Urinary system disorders	109 (0.1)	0 (0.0)	2 (1.8)	107 (98.2)	0 (0.0)	100 (91.7)
Bulimia nervosa	76 (0.1)	54 (71.1)	0 (0.0)	12 (15.8)	10 (13.2)	22 (29.0)
Other	317 (0.3)	61 (19.2)	11 (3.5)	126 (39.8)	119 (37.5)	211 (66.6)
ANY INDICATION	101,759 (100.0)	43,462 (42.7)	23,898 (23.5)	10,936 (10.8)	23,463 (23.1)	29,907 (29.4)

Abbreviations: SSRI = selective serotonin reuptake inhibitor, SNRI = serotonin-norepinephrine reuptake inhibitor, TCA = tricyclic antidepressant

[a]1.8% of prescriptions had multiple treatment indications recorded and were assigned to multiple categories. As a result, the sum of prescriptions across the individual treatment indication categories exceeds the number of prescriptions for any indication (last row).

[b]Percentages were calculated using the total number of antidepressant prescriptions for any indication (n=101,759) as the denominator.

[c]Percentages for each pharmacological class were calculated using the total number of prescriptions for the indication as the denominator.

[d]Includes citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline.

[e]Includes desvenlafaxine, duloxetine, and venlafaxine.

[f]Includes amitriptyline, clomipramine, desipramine, doxepin, imipramine, nortriptyline, and trimipramine.

[g]Includes bupropion, maprotiline, mirtazapine, trazodone, and vortioxetine.

[h]For each treatment indication category, a prescription was classified as off-label if the drug was not approved for the indication by Health Canada or the FDA as of September 2015. For any indication (last row), a prescription was classified as off-label if the drug was not approved for all of its recorded indications. Percentages were calculated using the total number of prescriptions for the indication as the denominator.

[i]Includes anxiety, generalized anxiety disorder, and other anxiety disorders except panic disorders and phobias.

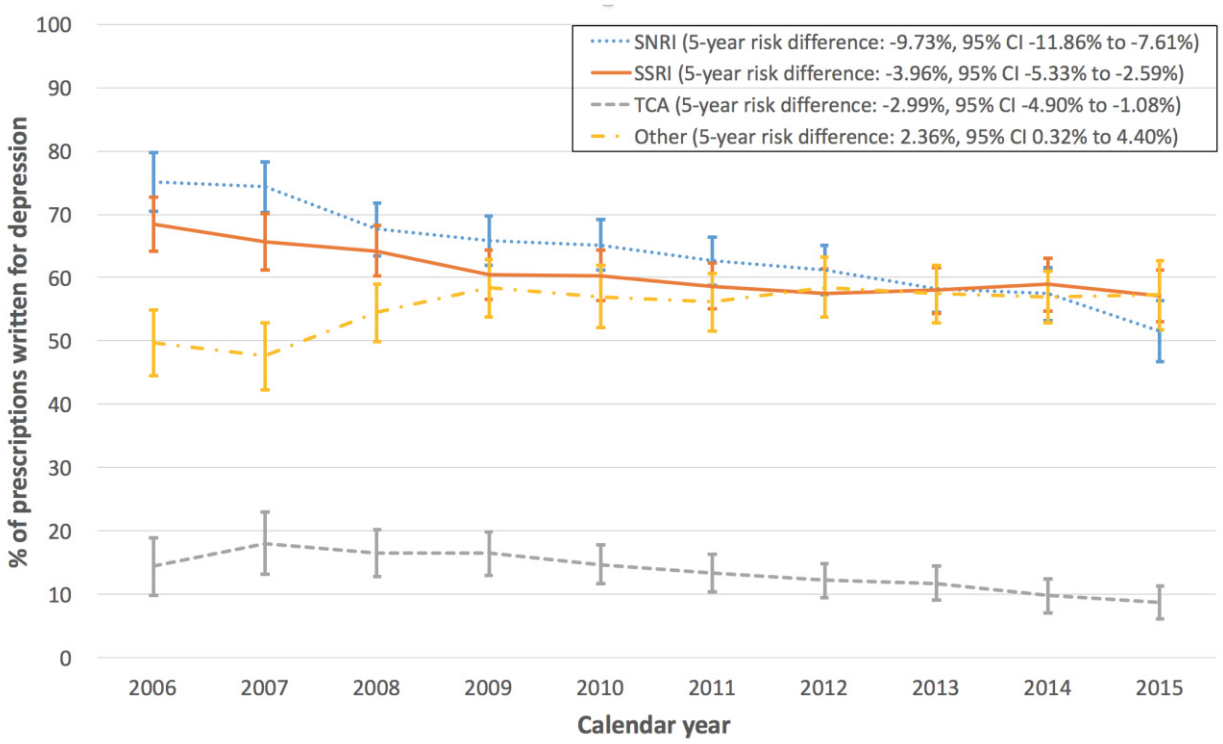


Figure 4-1. Percentage of antidepressant prescriptions for depression by pharmaceutical class, 2006-2015

Abbreviations: SNRI = serotonin-norepinephrine reuptake inhibitor, SSRI = selective serotonin reuptake inhibitors, TCA = tricyclic antidepressants.

The plots show the unadjusted percentage of antidepressant prescriptions written for depression in each calendar year, by pharmacological class. The error bars represent 95% CIs that were calculated based on standard errors corrected for multi-level clustering of prescriptions using an alternating logistic regression algorithm. Five-year risk difference estimates in the percentage of antidepressant prescriptions for depression were obtained from a generalized linear risk difference model (with an identity link) that included a linear effect of calendar time and dummy variables for individual pharmaceutical classes, along with their interactions with calendar time. All risk difference estimates were adjusted for patient age and sex and used an alternating logistic regression algorithm to account for multi-level clustering of prescriptions. As there were no missing data on patient age or sex, all prescriptions were included in the regression model.

5 Off-label indications for antidepressants in primary care: descriptive study of prescriptions from an indication-based electronic prescribing system

5.1 Preamble

In manuscript 1, I found that nearly one-third of all antidepressant prescriptions in primary care were written for an off-label indication. However, off-label drug use is only a pharmacovigilance concern when it lacks strong scientific evidence (88). Thus, in this next manuscript, I took a closer look at off-label indications for antidepressant prescriptions and examined the level of scientific evidence supporting these uses. I also investigated the extent to which intra-class substitution (i.e. prescribing an unlicensed drug for an indication instead of another licensed drug in the same class) may contribute to off-label prescribing.

The results from this manuscript provide interesting insights into the patterns of antidepressant prescribing for off-label indications and further motivate the need for increased pharmacovigilance around antidepressants.

This manuscript was published in *The BMJ* in February 2017. Appendix C contains a reprint of the original article.

5.2 Title page and footnotes

Title: Off-label indications for antidepressants in primary care: descriptive study of prescriptions from an indication based electronic prescribing system

Jenna Wong[1], MSc; Aude Motulsky[2], PhD; Tewodros Eguale[3], MD, PhD; David L Buckeridge[1], MD, PhD; Michal Abrahamowicz[1], PhD; and Robyn Tamblyn[1], PhD

Affiliations:

[1]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

[2] Department of Management, Evaluation and Health Policy, School of Public Health, Université de Montréal

[3]School of Pharmacy, Massachusetts College of Pharmacy and Health Sciences (MCPHS), Boston, MA

Corresponding author:

Jenna Wong

1140 Pine Avenue West

Montreal, Quebec, Canada

H3A 1A3

Email: jenna.wong@mail.mcgill.ca

Citation: 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. *The BMJ*. 2017;356:j603.

5.3 Abstract

Objective: To examine off-label indications for antidepressants in primary care and determine the level of scientific support for off-label prescribing.

Design: Descriptive study of antidepressant prescriptions written by primary care physicians using an indication-based e-prescribing system.

Setting: Primary care practices around two major urban centers in Quebec, Canada.

Participants: Patients aged 18 years or older who visited a study physician between 1 January 2003 and 30 September 2015 and had an antidepressant prescribed using the e-prescribing system.

Study measurements: The prevalence of off-label indications for antidepressant prescriptions by class and by individual drug. Among off-label antidepressant prescriptions, the proportion of prescriptions where the indication had either strong scientific evidence for the prescribed drug, no strong evidence for the prescribed drug but strong evidence for another drug in the same class, or no strong evidence for the prescribed drug and all other drugs in the same class.

Results: 106,850 antidepressant prescriptions were written by 174 physicians for 20,920 adults. By class, tricyclic antidepressants had the highest prevalence of off-label indications (81.4%; 95% CI, 77.3% - 85.5%) largely due to a high off-label prescribing rate for amitriptyline (93%; 95% CI, 89.6% - 95.7%). The use of trazodone for insomnia was the most common off-label use for antidepressants, accounting for 26.2% (95% CI, 21.9% - 30.4%) of all off-label prescriptions. Only 15.9% (95% CI, 13.3% - 18.6%) of all off-label prescriptions were supported by strong scientific evidence, yet for 39.6% (95% CI, 35.7% - 43.2%), there was another antidepressant in the same class with strong evidence for the respective indication. For the remaining 44.6% (95% CI, 40.2% - 49.0%) of off-label prescriptions, neither the prescribed drug nor any other drugs in the class had strong evidence for the indication.

Conclusions: When primary care physicians prescribed antidepressants for off-label indications, these indications were usually not supported by strong scientific evidence, yet often another antidepressant in the same class existed that had strong evidence for the respective indication. There is an important need to generate and provide physicians with evidence on off-label antidepressant use to optimize prescribing decisions.

5.4 Introduction

Antidepressant use has increased substantially in the United Kingdom (89,90) and other Western countries such as Canada (12) and the United States (86). In fact, the number of antidepressants dispensed in England increased by 3.9 million (6.8%) between 2014 and 2015 – more than any other therapeutic class of prescription drugs (90). One suspected factor underlying the widespread use of antidepressants is an expanding array of indications for these drugs, many of which are unapproved (off-label) for certain antidepressants (15).

There is a lack of epidemiological evidence on the extent to which physicians prescribe antidepressants for off-label indications because treatment indications are not documented for most prescriptions (29). With the advent of e-prescribing systems, however, formal documentation of treatment indications linked to prescriptions (i.e. indication-based prescribing) is possible. Although indication-based prescribing is not broadly used at the moment, it represents a valuable means for studying off-label prescribing (9). We recently used data from a unique indication-based e-prescribing system to describe treatment indications for antidepressants in primary care (91). We found that over the last decade, primary care physicians commonly and increasingly prescribed antidepressants for non-depressive indications. Moreover, when antidepressants were not prescribed for depression, two out of three prescriptions were for an off-label indication.

Off-label prescribing warrants particular attention and oversight when the medication use is not supported by scientific evidence showing greater benefits relative to risk (92,93). Inefficacious antidepressant use is a concern because it creates unnecessary costs and puts patients at risk of experiencing burdensome side effects and serious adverse events that could be avoided. For example, even though newer generation antidepressants like selective serotonin reuptake inhibitors (SSRIs) are considered safer and more tolerable than the older generation tricyclic antidepressants (TCAs), they are costly and have been associated with notable side effects (e.g. sexual dysfunction, drowsiness, insomnia, weight gain, and fatigue

(24,25,94)) and safety concerns including an increased risk of fractures (95) and upper gastrointestinal bleeds (96,97). Off-label antidepressant use could also expose patients to unknown health risks if their clinical characteristics differ from the patient populations studied in pre-market clinical trials (98). Indeed, the risk of adverse drug events has been found to be 54% higher when drugs are used off-label without strong scientific evidence than when drugs are used on-label (99).

Although an estimated 29% of antidepressants are prescribed for off-label indications(91), it is unknown to what extent these off-label prescriptions are supported by scientific evidence. Thus, the objective of this study was to examine off-label indications for antidepressants in primary care and assess the level of scientific evidence supporting these off-label prescriptions.

5.5 Methods

5.5.1 Study design and setting

This was a descriptive study that took place in the Canadian province of Quebec, where a universal health insurance program covers the cost of essential medical care for all residents. By law, all residents must be covered for prescription drugs through either private plans (i.e. group or employee benefit plans) or the public drug insurance plan. Approximately 50% of residents are registered in the public drug insurance plan, including residents >65 years old, welfare recipients, and persons not insured through an employer. At a minimum, all private plans must provide the same formulary for insured drugs as the public drug insurance plan (100).

5.5.2 Data source and study population

The Medical Office of the XXIst Century (MOXXI) is an electronic prescription and drug management system used by consenting primary care physicians in community-based, fee-for-service practices around two major urban centers in Quebec(36). 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 the MOXXI program and have their information used for research purposes.

The e-prescribing tool in the MOXXI system requires physicians to explicitly document at least one treatment indication per prescription by either using a drop-down menu that lists on-label and off-label indications (without distinction) or typing the indication(s) into a free-text field. In a validation study(37), these physician-documented indications had excellent sensitivity (98.5%) and high positive predictive value (97.0%) when compared to a blinded, post-hoc physician-facilitated chart review. The MOXXI system also provides physicians with access to professional drug monographs that are maintained by a commercial vendor (83) and produces automated drug alerts about potential prescribing problems. Alerts are generated when potential dosing errors or drug-drug, drug-disease, drug-age, or drug-allergy contraindications are identified; however, alerts are not generated when drugs are prescribed for off-label indications.

This study was approved by the McGill Institutional Review Board.

5.5.3 Inclusion and exclusion criteria

This study included prescriptions of drugs approved for depression that were written by MOXXI physicians between 1 January 2003 and 30 September 2015 for patients aged 18 years or older. The antidepressant prescription was the unit of analysis. Drugs with fewer than 150 prescriptions during the study period (roughly corresponding to a prescribing frequency of less than once per month) were excluded, which resulted in the exclusion of all monoamine oxidase inhibitors (phenelzine, tranylcypromine, moclobemide, and isocarboxazid), nefazodone, maprotiline, and vortioxetine.

5.5.4 Measurements

5.5.4.1 On-label versus off-label indications

Treatment indications were first categorized based on the International Classification of Diseases and Related Health Problems, 10th Revision. Each prescription – representing a drug-indication pair – was then classified as on-label or off-label depending on whether the drug had been approved for the indication by Health Canada or the US Food and Drug Administration as

of September 2015 (the end of the study period). Approved indications were determined at the end of the study period rather than the year in which the prescription was written so that all prescriptions would be classified using the same benchmark. If a physician recorded multiple indications for the drug (n=1,922 or 1.8% of all antidepressant prescriptions), the prescription was classified as off-label only if all the indications were not approved.

5.5.4.2 Level of scientific evidence for off-label prescriptions

Off-label prescriptions were further analyzed according to the level of scientific evidence supporting the drug's use for the off-label indication. Off-label prescriptions were assigned to one of three categories: 1) strong evidence for the prescribed drug, 2) no strong evidence for the prescribed drug but strong evidence for another drug in the same class, or 3) no strong evidence for the prescribed drug and all other drugs in the same class. To determine if off-label prescriptions had strong evidence for the prescribed drug, we used the DRUGDEX compendium (Thomson Micromedex) (84), which is a reputable and authoritative reference used by the US Centers for Medicare and Medicaid Services to determine coverage for off-label drug uses (101).The compendium contains evaluations of drug efficacy, strength of recommendation, and strength of evidence for off-label drug-indication pairs. Using the same criteria as in previous studies(9,99,102), off-label prescriptions were classified as having strong evidence for the prescribed drug if evidence showed that the drug was effective or favoured efficacy for the indication, the drug was recommended for all or most patients with the indication, and at least one randomized clinical trial was included among the studies used to evaluate the drug's efficacy for the indication. If an off-label prescription did not have strong evidence for the prescribed drug, we then determined if there was strong evidence for another drug in the same class. This condition was satisfied if another drug in the same class was either on-label or off-label with strong evidence for the indication. If an off-label prescription still did not have strong evidence for another drug in the class, then the prescription was classified as having no strong evidence for the prescribed drug and all other drugs in the same class.

5.5.5 Statistical analysis

Patient and physician characteristics were summarized using descriptive statistics. The prevalence of off-label indications was estimated as the number of off-label prescriptions divided by the total number of antidepressant prescriptions overall, in the class, or for the individual drug. The level of scientific evidence for off-label prescriptions was estimated as the number of off-label prescriptions in each evidence category divided by the total number of off-label antidepressant prescriptions overall or in the class. The prevalence of different treatment indications for each drug was estimated as a proportion using the total number of prescriptions for the drug as the denominator. For all proportions, 95% confidence intervals (CI) were calculated using a cluster bootstrap approach (103) to account for within-cluster correlation among prescriptions for the same patient and from the same physician. The reported 95% CIs correspond to the values of the 2.5^{th} and 97.5^{th} percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples (103). All analyses were conducted using SAS (SAS Institute) software, version 9.4.

5.5.6 Patient involvement

No patients were involved in setting the research question or the study measures, nor were they involved in developing plans for the design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. The findings from this study will be disseminated to study participants though physician newsletters and patient-friendly handouts.

5.6 Results

During the study period, 106,850 antidepressant prescriptions (5.8% of 1.83 million prescriptions for any drug) were written by 174 primary care physicians for 20,920 adults. There were approximately equal numbers of male (n=90; 51.7%) and female (n=84; 48.3%) physicians, most of whom had been trained in North America (n=160; 92.0%) and practicing for at least 15 years (n=131; 75.3%). Two-thirds of patients were female (n=13,990; 66.9%), most patients were middle-aged at the time of their earliest antidepressant prescription (median of 53 years,

interquartile range [IQR] 43-65 years), and patients were equally likely to have public (n=10,875; 52.0%) or private (n=10,045; 48.0%) drug insurance. Over the study period, patients had a median of 3 (IQR 1 – 7) antidepressant prescriptions and were prescribed a median of 1 (IQR 1 – 2) type of antidepressant drug.

5.6.1 Prevalence of off-label indications

Overall, 29.3% (95% CI, 26.6% to 32.3%) of all antidepressant prescriptions were written for an off-label indication (Table 5-1). By class, TCAs had the highest prevalence of off-label indications (81.4%; 95% CI, 77.3% to 85.5%), followed by 'other' antidepressants (42.4%; 95% CI, 37.1% to 47.7%) and SSRIs (21.8%; 95% CI, 19.0% to 25.0%). In contrast, the prevalence of off-label indications was much lower for serotonin-norepinephrine reuptake inhibitors (SNRIs) (6.1%; 95% CI 4.8% to 7.5%). The high prevalence of off-label indications for TCAs was mostly due to amitriptyline, which was only approved for depression but was almost exclusively prescribed for off-label indications (93.0%; 95% CI, 89.6% to 95.7%) – most commonly pain (48.4%; 95% CI, 39.7% to 57.8%), insomnia (22.5%; 95% CI, 13.6% to 31.3%), and migraine (16.7%; 95% CI, 12.2% to 21.9%) (Table 5-2). The high prevalence of off-label indications among 'other' antidepressants was largely due to trazodone, which was mostly prescribed for insomnia (82.5%; 95% CI, 74.5% to 88.1%) even though it was not approved for this indication. SSRIs and SNRIs had a lower prevalence of off-label indications because compared to TCAs, they were more frequently prescribed for depression, which by definition was an approved indication for all antidepressants (Table 5-2).

5.6.2 Level of scientific evidence for off-label indications

Among all off-label antidepressant prescriptions, there were 143 unique drug-indication pairs – the most common of which were trazodone for insomnia (representing 26.2%; 95% CI, 21.9% to 30.4% of all off-label prescriptions), citalopram for anxiety (17.8%; 95% CI, 14.8% to 21.3%), amitriptyline for pain (13.8%; 95% CI, 11.0% to 16.9%), and amitriptyline for insomnia (6.4%; 95% CI, 3.9% to 9.5%) (data not shown). Only three of these 143 off-label drug-indication pairs met the pre-defined criteria (9,99,102) for having strong scientific evidence: amitriptyline (a

TCA) for pain, escitalopram (an SSRI) for panic disorders, and venlafaxine (an SNRI) for obsessive-compulsive disorder. These three pairs collectively comprised 15.9% (95% CI, 13.0% to 19.3%) of all off-label antidepressant prescriptions (Table 5-1) – the majority of which were amitriptyline prescriptions for pain (representing 87.1%; 95% CI, 80.9% to 92.1% of all off-label prescriptions with strong evidence for the prescribed drug). As a result, the proportion of off-label antidepressant prescriptions with strong evidence for the prescribed drug was much higher for TCAs (45.7%; 95% CI, 37.8% to 54.0%) compared to SNRIs (11.0%; 95% CI, 4.6% to 18.4%) and SSRIs (4.7%; 95% CI, 2.7% to 7.2%) (Table 5-1).

Off-label antidepressant prescriptions had strong evidence for another drug in the same class – but not the prescribed drug – in 39.6% (95% CI, 35.7% to 43.2%) of all cases (Table 5-1). This proportion was highest among off-label SSRI prescriptions (92.0%; 95% CI, 89.2% to 94.4%), while lower among off-label prescriptions for SNRIs (35.4%; 95% CI, 25.0% to 46.7%) and TCAs (28.3%; 95% CI, 20.5% to 36.6%). This proportion was not assessed for 'other' antidepressants since trazodone, bupropion, and mirtazapine were not considered as part of the same class.

For the remaining 44.5% (95% CI, 40.2% to 49.0%) of off-label antidepressant prescriptions, neither the prescribed drug nor any other drug in the same class had strong evidence for the indication (Table 5-1). All off-label prescriptions for 'other' antidepressants were classified in this evidence category. The proportion of off-label prescriptions with no scientific support for any drug in the class was also quite high for SNRIs (53.7%, 95% CI, 40.6% to 66.6%) and TCAs (26.0%; 95% CI, 21.2% to 31.1%), but was much lower for SSRIs (3.3%; 95% CI, 2.0% to 4.8%).

5.7 Discussion

This is the first study to provide evidence on the level of scientific support for off-label antidepressants prescriptions, the prevalence of off-label indications for individual antidepressants, and the most common off-label uses for antidepressants. Nearly one-third (29%) of all antidepressants in this study were prescribed for an off-label indication, as found previously (91). Among all off-label antidepressant prescriptions, only one in six prescriptions

was supported by strong scientific evidence, but there was often another antidepressant in the same class with strong evidence that could have been considered instead, especially among off-label SSRI prescriptions. Still, nearly half of all off-label antidepressant prescriptions did not have strong evidence for the prescribed drug and all other antidepressants in the same class. Among the many off-label uses for antidepressants, physicians most frequently prescribed trazodone for insomnia even though this use was not evidence-based.

5.7.1 Comparison with other studies

Few published studies exist on off-label prescribing due to challenges associated with measuring diagnoses (indications) for prescriptions. Compared to our findings where 29% of antidepressant prescriptions were off-label, Chen *et al.* (34) found that 75% of Georgia Medicaid enrollees who received antidepressants received at least one antidepressant off-label. The rate of off-label antidepressant use was notably higher in this study because the authors classified prescriptions as off-label if the patient did not have a diagnostic code for an approved indication recorded in administrative claims data during the same year. This methodology most likely overestimated the off-label prescribing rate since diagnostic codes in administrative data are often incomplete or inaccurate, especially for psychiatric conditions (104).

Only three studies – one Canadian (9) and two US (10,102) – have used documented treatment indications to study off-label prescribing, none of which focused specifically on antidepressants. Eguale *et al.* (9) combined antidepressants with other central nervous system drugs but reported fairly comparable results, with 26% of prescriptions for off-label indications – 18% of which were supported by strong evidence. Radley *et al.* (10) combined antidepressants with anxiolytics and antipsychotics, but again reported a similar off-label prescribing rate of 31%. However, the proportion of off-label prescriptions with strong scientific support in this study was notably lower than ours at only 6%, possibly due to the inclusion of other psychiatric drugs or because evidence to support some off-label antidepressant uses had not been generated at the time of the analysis. Finally, Walton *et al.* (102) presented results for only five antidepressants but similarly found that amitriptyline and trazodone were the antidepressants

most frequently prescribed for off-label indications. However, their off-label prescribing rate was notably lower for amitriptyline (69%) and trazodone (43%), possibly reflecting inter-country differences in the use of antidepressants versus other medications to treat pain and insomnia.

In all of these studies, none of the authors assessed the proportion of off-label antidepressant prescriptions where the prescribed drug did not have strong evidence but another antidepressant from the same class existed that had strong evidence for the respective indication.

5.7.2 Potential explanations for off-label prescribing

Several contextual factors may contribute to physicians prescribing antidepressants for off-label indications. First, the vast and increasing number of drugs on the market makes it challenging for physicians to keep track of which indications are approved for specific products (105), especially when pharmaceutical companies have been known to promote drug use for off-label indications (106). Second, constraints such as the list of drugs included on patients' health plan formularies may influence which drugs physicians prescribe, especially if physicians presume that drugs in the same class are interchangeable (107,108). For example, in our setting, escitalopram was not covered for patients enrolled in the public drug insurance plan. Interestingly, we found that when study physicians prescribed SSRIs to patients with public drug insurance, they infrequently prescribed escitalopram (4.7% of all SSRI prescriptions for patients with public drug insurance) but frequently prescribed citalopram (51.4%). However, for patients with private drug insurance, study physicians equally prescribed escitalopram and citalopram (29.3% and 31.7% of all SSRI prescriptions for patients with private drug insurance, respectively). Third, primary care physicians may prescribe antidepressants off-label because alternative therapies for a given indication are contraindicated or perceived as higher risk medications. For example, benzodiazepines and Z-drugs like zolpidem and zaleplon have been shown to be efficacious for treating insomnia (109). However, these drugs have been labelled as potentially inappropriate medications for older adults, and if prescribed, could even negatively affect providers' quality and performance measures (110). Many physicians who are

concerned about the health of their older patients may consequently prescribe trazodone instead because they believe it is a safer medication. Finally, many off-label indications for antidepressants are symptom-based conditions for which few approved medications exist. Primary care physicians may struggle to find effective treatments for these conditions and prescribe antidepressants as a last resort, indicating a gap in needed pharmacotherapy.

5.7.3 Implications of findings

For both primary care physicians and specialists (since specialists may initiate antidepressant therapy that is then continued by a primary care physician), our findings emphasize the importance of considering the level of evidence supporting risk-benefit when prescribing an antidepressant, especially if the drug is known to have important adverse side effects (111). When evidence to support efficacy is lacking, physicians should exercise caution, prescribe conservatively, and inform patients of this information via a shared decision-making process (111). This ideal, however, is challenging to achieve because physicians face time constraints, both the drug market and scientific literature are vast and ever-evolving, and many physicians find it challenging to critically appraise and interpret the results of epidemiological studies (112). Indication-based e-prescribing systems that are integrated with clinical decision support tools could help overcome these obstacles by notifying physicians when drugs are being prescribed off-label without supporting evidence and providing them with access to concise, up-to-date summaries of the available evidence. Providing the public with access to patient-friendly resources about the level of scientific evidence supporting different treatment options for a given indication could also further facilitate the decision-making process between physicians and patients.

Our finding that among off-label prescriptions, 40% were for indications where the prescribed drug did not have strong evidence but another drug in the same class was approved or supported by strong evidence is clinically important. Many physicians may view this type of off-label prescribing as different from off-label prescribing without scientific evidence for the entire class because they assume that drugs within the same class are interchangeable (113,114).

However, “class effects” cannot be assumed because even slight differences in chemical structure between drugs can alter their pharmacodynamic and pharmacokinetic properties, leading to clinically relevant differences in efficacy and risk.(114) For example, statins have been shown to differ not only in efficacy (115) but also in safety, as demonstrated by the withdrawal of cerivastatin from the market in 2001 because the risk of rhabdomyolysis was 10 times higher for cerivastatin than other statins (116). Clinical guidelines recommend that when physicians select a particular drug to prescribe, they should consider the level of scientific evidence supporting the specific drug (117). It should not be assumed that all drugs within a class are likely to be efficacious for treating an indication when one member of the class has proven efficacy.

Finally, more evidence is needed on the clinical outcomes associated with antidepressant use for off-label indications. However, within a context of limited resources, it is unlikely that randomized clinical trials will be conducted for each off-label drug-indication pair, especially for older drugs that are no longer owned by an innovator company (92). Thus, in addition to randomized clinical trials, post-market drug surveillance systems represent valuable resources for assessing off-label antidepressant use. Such systems face challenges associated with measuring treatment indications and patient-reported outcomes, but these challenges could be overcome by increasing the use of indication-based e-prescribing systems and electronic health records that track patient outcomes. Indeed, this study demonstrates the benefits that indication-based prescribing can have towards addressing knowledge gaps around off-label antidepressant prescribing.

5.7.4 Strengths and limitations

The key strength of this study is that it included more than 12 years of antidepressant prescriptions from an e-prescribing system where physicians systematically documented treatment indications at the point of prescribing. However, study participants were from one Canadian province where prescribers were generally younger and patients were generally older with more health complexities (80). These characteristics may influence the generalizability of

our findings since younger physicians are more likely to prescribe drugs off-label without scientific evidence, while sicker patients are less likely to receive off-label prescriptions (9). Another study strength is that physicians were unlikely to have altered their true responses when recording indications in the e-prescribing system because the drop-down menu did not distinguish between on-label and off-label indications for a drug. On the other hand, we could not identify when physicians consciously prescribed antidepressants off-label. Indeed, a portion of antidepressants in this study may have been prescribed off-label for a specific reason (e.g. patient experienced side effects to another drug in the same class, formulary restrictions, etc.).

5.7.5 Study considerations

First, our estimates of off-label antidepressant prescribing were conservative because we did not consider other aspects of off-label drug use (e.g. dose, frequency, duration of treatment) and we used the approved indications and available evidence at the *end* of the study period. Second, we presumed that approved indications for drugs were backed by strong scientific evidence, which may not have been true in some cases given that the quality of clinical trial evidence used by regulatory agencies as the basis for approving new therapeutics and supplemental indications has been shown to vary widely (118,119). Third, to identify evidence-based off-label uses for antidepressants, we used pre-established criteria that has been used in other studies (9,99,102). However, our list of evidence-based antidepressants for each indication may not always be identical to the recommendations from clinical guidelines. For example, recommendations from two national guidelines for managing anxiety-related disorders (117,120) are similar but slightly more inclusive than ours. Finally, because regulatory bodies in North America and Europe are not entirely harmonized in their list of approved indications for drugs (e.g. amitriptyline is approved for chronic pain therapy in Europe but not in North America), rates of off-label antidepressant use may naturally be higher in North America than in Europe.

5.8 Conclusions

By using information from an indication-based e-prescribing system, we found that when primary care physicians prescribed antidepressants for off-label indications, the prescribed drug was usually not supported by strong evidence for the respective indication. However, there was often another drug in the same class with strong evidence that could have been considered. These findings highlight an urgent need to produce more evidence on the risks and benefits of off-label antidepressant use and to provide physicians with this evidence at the point of prescribing. Technologies such as indication-based e-prescribing systems and electronic health records have the potential to become essential components of effective post-market drug surveillance systems for monitoring and evaluating off-label antidepressant use. By integrating these technologies with knowledge databases and clinical decision support tools, they could also provide an effective means for communicating evidence back to physicians to optimize prescribing decisions.

Table 5-1. Proportion of antidepressants prescribed for off-label indications and their level of evidence, by pharmacological class

Pharmacological class	No. of prescriptions	Off-label indication		LEVEL OF EVIDENCE FOR OFF-LABEL INDICATIONS					
				Strong evidence for the prescribed drug[d]		No strong evidence for the prescribed drug, but strong evidence for another drug in the same class[e]		No strong evidence for the prescribed drug and all other drugs in the same class	
		N	%[a] (95% CI[c])	N	%[b] (95% CI[c])	N	%[b] (95% CI[c])	N	%[b] (95% CI[c])
SSRI	45,608	9,960	21.8 (19.0 to 25.0)	473	4.7 (2.7 to 7.2)	9,160	92.0 (89.2 to 94.4)	327	3.3 (2.0 to 4.8)
SNRI	25,235	1,539	6.1 (4.8 to 7.5)	169	11.0 (4.6 to 18.4)	544	35.4 (25.0 to 46.7)	826	53.7 (40.6 to 66.6)
TCA	11,645	9,480	81.4 (77.3 to 85.5)	4,335	45.7 (37.8 to 54.0)	2,682	28.3 (20.5 to 36.6)	2,463	26.0 (21.2 to 31.1)
Other[f]	24,362	10,340	42.4 (37.1 to 47.7)	0	0.0 (0.0 to 0.0)	N/A[g]	N/A[g]	10,340	100.0 (100.0 to 100.0)
All classes	**106,850**	**31,319**	**29.3 (26.6 to 32.3)**	**4,977**	**15.9 (13.0 to 19.3)**	**12,386**	**39.6 (35.7 to 43.2)**	**13,956**	**44.6 (40.2 to 49.0)**

Abbreviations: CI = confidence interval; SSRI = selective serotonin reuptake inhibitor; SNRI = serotonin-norepinephrine reuptake inhibitor; TCA = tricyclic antidepressant.

[a]Percentages were calculated using the total number of prescriptions in the class as the denominator.

[b]Percentages were calculated using the number of prescriptions in the class that were written for an off-label indication as the denominator.

[c]Confidence intervals were calculating using a cluster bootstrap approach (103) to account for non-independence of prescriptions from the same physician and patient. The reported 95% CIs correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples.

[d]Based on evaluations from the DRUGDEX compendium in three dimensions: efficacy, strength of recommendation, and strength of evidence. Prescriptions for an off-label indication were classified as having strong evidence for the prescribed drug if they met three criteria: 1) evidence showed that the drug was effective or favored efficacy for the indication, 2) the drug was recommended for all or most patients with the indication, and 3) at least one randomized controlled trial was included among the studies used to evaluate the drug's efficacy for the indication.

[e]Defined as off-label prescriptions where the prescribed drug did not have strong evidence for the indication, but another drug in the class was either on-label or off-label with strong evidence for the indication based on evaluations from the DRUGDEX compendium.

[f]Includes trazodone, bupropion, and mirtazapine.

[g]Not assessed for drugs in this category because they were not considered as part of the same class.

Table 5-2. Off-label indications and most common antidepressant treatment indications, by drug

Drug name, by pharmacological class	Total N	Off-label indication		TREATMENT INDICATIONS[a]								
		N	%[b] (95% CI[c])	Most common			Second most common			Third most common		
				Indication	N	%[b] (95% CI[c])	Indication	N	%[b] (95% CI[c])	Indication	N	%[b] (95% CI[c])
SSRI												
Citalopram	19,480	6,988	35.9 (31.5 to 40.9)	**Depression**	12,492	64.1 (59.1 to 68.5)	Anxiety disorder[d]	5,745	29.5 (25.0 to 34.6)	Panic disorder	882	4.5 (2.7 to 7.2)
Paroxetine	9,212	94	1.0 (0.4 to 1.9)	**Depression**	4,476	48.6 (40.2 to 57.3)	**Anxiety disorder[d]**	2,719	29.5 (23.6 to 36.0)	**Panic disorder**	1,563	17.0 (10.6 to 23.8)
Escitalopram	7,108	601	8.5 (5.7 to 11.4)	**Depression**	4,354	61.3 (55.2 to 67.0)	**Anxiety disorder[d]**	2,075	29.2 (23.2 to 35.3)	Panic disorder	503	7.1 (4.6 to 9.9)
Sertraline	6,805	1,680	24.7 (18.6 to 31.8)	**Depression**	4,383	64.4 (55.4 to 71.8)	Anxiety disorder[d]	1,847	27.1 (20.6 to 34.1)	**Panic disorder**	398	5.8 (2.4 to 10.0)
Fluoxetine	2,079	322	16.0 (10.0 to 24.5)	**Depression**	1,566	75.3 (65.1 to 83.2)	Anxiety disorder[d]	249	12.0 (7.1 to 18.2)	**Panic disorder**	64	3.1 (0.2 to 7.4)
Fluvoxamine	924	265	28.7 (15.6 to 45.3)	**Depression**	592	64.1 (47.5 to 76.5)	Anxiety disorder[d]	233	25.2 (13.8 to 41.9)	**OCD**	71	7.7 (2.9 to 15.5)
SNRI												
Venlafaxine	21,369	1,501	7.0 (5.5 to 8.7)	**Depression**	14,282	66.8 (62.3 to 71.2)	**Anxiety disorder[d]**	5,053	23.6 (19.4 to 27.9)	**Panic disorder**	782	3.7 (2.6 to 4.9)
Duloxetine	2,969	9	0.3 (0.0 to 1.0)	**Depression**	1,139	38.4 (31.4 to 45.4)	**Pain**	1,053	35.5 (27.8 to 43.1)	**Fibromyalgia**	604	20.3 (14.5 to 27.1)
Desvenlafaxine	897	29	3.2 (0.7 to8.0)	**Depression**	868	96.8 (91.9 to 99.3)	Anxiety disorder[d]	16	1.8 (0.0 to 5.5)	Menopausal hot flashes	6	0.7 (0.0 to 1.8)

Abbreviations: CI = confidence interval; SSRI = selective serotonin reuptake inhibitor; SNRI = serotonin-norepinephrine reuptake inhibitor; OCD = obsessive-compulsive disorder; ADHD = attention deficit/hyperactivity disorder.

[a]Indications in **bold** were approved for the drug by Health Canada or the US Food and Drug Administrations as of September 2015 (the end of the study period).

[b]Percentages were calculated using the total number of prescriptions for the drug as the denominator.

[c]Confidence intervals were calculated using a cluster bootstrap approach(103) to account for non-independence of prescriptions from the same physician and patient. The reported 95% CIs correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples(103).

[d]Includes anxiety, generalized anxiety disorder and other anxiety disorder. Excludes panic disorder, phobias, obsessive-compulsive disorder, and post-traumatic stress disorder.

Table 5-2. (continued) Off-label indications and most common antidepressant treatment indications, by drug

Drug name, by pharmacological class	Total N	Off-label indication		TREATMENT INDICATIONS[a]								
		N	%[b] (95% CI[c])	Most common			Second most common			Third most common		
				Indication	N	%[b] (95% CI[c])	Indication	N	%[b] (95% CI[c])	Indication	N	%[b] (95% CI[c])
TCA												
Amitriptyline	8,993	8,361	93.0 (89.6 to 95.7)	Pain	4,349	48.4 (39.7 to 57.8)	Insomnia	2,023	22.5 (13.6 to 31.3)	Migraine	1,501	16.7 (12.2 to 21.9)
Doxepin	782	92	11.8 (3.2 to 21.3)	**Insomnia**	285	36.4 (24.1 to 49.3)	**Depression**	171	21.9 (9.7 to 35.9)	**Anxiety disorder[d]**	150	19.2 (9.0 to 32.0)
Nortriptyline	592	458	77.4 (59.9 to 89.5)	Pain	340	57.4 (35.0 to 74.1)	**Depression**	126	21.3 (9.5 to 38.5)	Anxiety disorder[d]	49	8.3 (2.4 to 20.0)
Trimipramine	562	165	29.4 (15.9 to 43.9)	**Depression**	397	70.6 (55.9 to 84.0)	Pain	93	16.5 (5.2 to 29.9)	Insomnia	22	3.9 (0.2 to 11.3)
Imipramine	285	218	76.5 (55.4 to 90.6)	Panic disorder	69	24.2 (1.6 to 42.5)	**Depression**	67	23.5 (9.4 to 44.5)	Urinary disorders	64	22.5 (4.7 to 54.6)
Desipramine	216	127	58.8 (30.5 to 80.7)	**Depression**	89	41.2 (19.2 to 69.1)	Pain	81	37.5 (12.0 to 61.4)	Anxiety disorder[d]	16	7.4 (1.1 to 21.3)
Clomipramine	215	59	27.4 (8.9 to 51.1)	**Depression**	107	49.8 (25.8 to 72.3)	**OCD**	49	22.8 (7.2 to 42.0)	Anxiety disorder[d]	36	16.7 (2.6 to 37.9)
Other												
Trazodone	10,070	8,938	88.8 (81.5 to 93.7)	Insomnia	8,303	82.5 (74.5 to 88.1)	**Depression**	1,132	11.2 (6.3 to 18.4)	Anxiety disorder[d]	574	5.7 (3.8 to 8.2)
Bupropion	8,384	780	9.3 (6.3 to 12.7)	**Depression**	7,052	84.1 (79.9 to 87.9)	**Nicotine dependence**	565	6.7 (4.4 to 9.6)	ADHD	372	4.4 (2.5 to 6.9)
Mirtazapine	5,908	622	10.5 (6.7 to 14.9)	**Depression**	5,286	89.5 (85.1 to 93.7)	Anxiety disorder[d]	473	8.0 (4.3 to 13.2)	Insomnia	157	2.7 (0.8 to 4.8)

Abbreviations: CI = confidence interval; TCA = tricyclic antidepressant; OCD = obsessive-compulsive disorder; ADHD = attention deficit/hyperactivity disorder.

[a]Indications in **bold** were approved for the drug by Health Canada or the US Food and Drug Administrations as of September 2015 (the end of the study period).

[b]Percentages were calculated using the total number of prescriptions for the drug as the denominator.

[c]Confidence intervals were calculated using a cluster bootstrap approach(103) to account for non-independence of prescriptions from the same physician and patient. The reported 95% CIs correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples(103).

[d]Includes anxiety, generalized anxiety disorder and other anxiety disorder. Excludes panic disorder, phobias, obsessive-compulsive disorder, and post-traumatic stress disorder.

6 Assessing the accuracy of using diagnostic codes from administrative data to infer antidepressant treatment indications: a validation study

6.1 Preamble

The findings from manuscripts 1 and 2 provided compelling evidence to indicate that more careful investigation is needed into the risks and benefits of off-label antidepressant use. However, the ability to conduct important pharmacovigilance activities for antidepressants is severely hampered by the fact that treatment indications for antidepressants are not usually documented. The literature review in Chapter 2 revealed that in the absence of documented treatment indications for antidepressants, many studies have used diagnostic codes to infer the medical reasons for antidepressant use. Although this approach is simple and feasible to use with large databases, diagnostic codes are often incomplete and inaccurate, especially for mental health conditions (35). Moreover, the accuracy of this approach has never been validated against a gold standard. Thus, the goal of the third manuscript was to address this knowledge gap by performing a validation study comparing the presence of diagnostic codes for plausible antidepressant treatment indications in administrative health data against the gold-standard treatment indications in the MOXXI system.

6.2 Title page and footnotes

Title: Assessing the accuracy of using diagnostic codes from administrative data to determine antidepressant treatment indications: a validation study

Jenna Wong[1], MSc; Michal Abrahamowicz[1], PhD; David L Buckeridge[1], MD, PhD; and Robyn Tamblyn[1], PhD

Affiliations:

[1]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

Corresponding author:

Jenna Wong

1140 Pine Avenue West

Montreal, Quebec, Canada

H3A 1A3

Email: jenna.wong@mail.mcgill.ca

6.3 Abstract

Background: Lack of information on treatment indications in administrative data presents a major obstacle for using these data to study antidepressant use for off-label indications. Several studies have used diagnostic codes from administrative data to infer antidepressant treatment indications, but this approach has never been validated against a reference standard.

Methods: This study analyzed antidepressant prescriptions written by primary care physicians in Quebec between 1 January 2003 and 31 December 2012 using an e-prescribing system that required physicians to record treatment indications. Prescriptions were linked to medical service claims and hospitalization data to obtain all recorded diagnostic codes. For 13 plausible conditions where antidepressants would be used, we determined whether patients had a diagnostic code for the condition recorded around the prescription date and compared this result to the physician-documented treatment indications in the e-prescribing system.

Results: The sensitivity of administrative diagnostic codes was very poor for all treatment indications, ranging from a high of only 31.2% (95% CI, 26.8% - 35.9%) for anxiety/stress disorders to as low as 1.3% (95% CI, 0.0% - 5.2%) for sexual dysfunction. Sensitivity was notably worse among older patients and patients with more chronic comorbidities. The positive predictive value of diagnostic codes varied widely between antidepressants of different therapeutic classes, where estimates were better among antidepressants that were more likely to be prescribed for the indication. Compared to hospitalization data, billings data were a better source of diagnostic codes for antidepressant treatment indications, with most of these codes being recorded by the prescribing physician.

Conclusions: Diagnostic codes from administrative data are poor proxies for antidepressant treatment indications and should not be used alone to infer treatment indications. Future work should determine whether other variables in administrative data besides diagnostic codes can improve the ability to predict antidepressant treatment indications.

6.4 Introduction

Nearly half of all antidepressants in primary care are prescribed for indications other than depression, including conditions such as insomnia, pain, and migraine (91). When antidepressants are not prescribed for depression, they are often prescribed for unapproved (off-label) indications, many of which are not evidence-based. In fact, we found that only 15.9% of all off-label antidepressant prescriptions were supported by strong scientific evidence (121). Given the frequency with which antidepressants are prescribed for non-evidence-based indications, post-market surveillance studies are needed to evaluate the safety and effectiveness of antidepressant use by indication.

The analysis of information from large administrative databases to monitor and evaluate the risks associated with indication-specific antidepressant use offers an attractive alternative to conducting clinical trials. Administrative databases can identify large, population-based cohorts of antidepressant users, capture real-life patterns of many different antidepressant uses, and detect rare outcomes or long-term effects that otherwise might not be observed in clinical trials (122). These databases also enable studies to be conducted in a relatively timely and cost-efficient manner (122). However, administrative databases do not contain information on treatment indications for drugs, which presents a major obstacle for using these data to evaluate indication-specific antidepressant use.

In the absence of documented treatment indications, a number of studies (30,31,34) have used diagnostic codes from administrative data to infer antidepressant treatment indications. This method has never been validated, so the potential biases introduced by this approach of inferring antidepressant treatment indications directly from diagnostic codes is not known. Thus, the objective of this study was to assess the accuracy of using diagnostic codes from administrative databases to determine antidepressant treatment indications, as compared to treatment indications recorded by the prescribing physician in an indication-based electronic prescribing (e-prescribing) system.

6.5 Methods

6.5.1 Context

This study took place in the Canadian province of Quebec, where a universal health insurance program covers the costs of essential medical care (including most services delivered by physicians and within hospitals) for all residents. Nearly all (>90%) physicians are reimbursed on a fee-for-service basis, with physicians submitting a claim to the provincial health insurance agency (the *Régie de l'assurance maladie du Québec* [RAMQ]) for each medical service provided in hospitals or private clinics (82). For each claim, physicians are requested – but not required – to provide a single diagnostic code representing the main reason for the visit (123). All diagnoses in billings data are recorded using the International Classification of Diseases, Ninth Revision (ICD-9) coding system. The province also maintains a hospitalization discharge summary database (MED-ECHO), which contains details of all hospitalizations occurring at acute care institutions in Quebec, including patient diagnoses. Each hospital discharge summary contains a principal diagnosis and up to 15 secondary diagnoses (82) (up to 25 secondary diagnoses starting in April 2006). These diagnoses are recorded by medical archivists based on chart review and were recorded using the ICD-9 system until April 2006 and the ICD-10 system after this date.

6.5.2 Study design

Thirteen different treatment indications for antidepressants were considered in this study, which included various on-label (83) and reported off-label indications (15,124–126). A separate validation study was conducted for each indication, where the unit of analysis was the prescription.

6.5.3 Data sources and inclusion criteria

The Medical Office of the XXIst Century (MOXXI) is an indication-based e-prescribing and drug management system used by consenting primary care physicians at community-based clinics around two major urban centers in Quebec (36). The e-prescribing tool in MOXXI requires physicians to document at least one treatment indication per prescription 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. In a previous study (37), these physician-documented treatment indications had excellent sensitivity (98.5%) and high positive predictive value (97.0%) when compared to a blinded, post-hoc physician-facilitated chart review. 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. In general, MOXXI physicians are more technologically proficient and have lighter patient loads with less fragmented care than non-MOXXI physicians, while MOXXI patients are older with more health complexities than non-MOXXI patients (80).

This study included all MOXXI prescriptions for any drug approved for depression written between 1 January 2003 and 31 December 2012. These prescriptions were linked to the RAMQ and MED-ECHO databases to get all diagnostic codes for patients recorded in physician billings or hospital discharge summary data within the past year. This study was approved by the McGill Institutional Review Board.

6.5.4 Study measurements

6.5.4.1 Antidepressant treatment indications

6.5.4.1.1 Reference standard

Antidepressant prescriptions were classified as positive for a given indication according to the reference standard (hereafter termed ‘reference positive’) if the prescribing physician documented the indication for the prescription in the MOXXI system. If the prescriber documented a subcategory of the indication under the ICD coding system (e.g. ‘panic attack’ as a subcategory of ‘anxiety disorders’), the prescription was also classified as reference positive for the indication. In cases where a prescription had multiple indications documented (only 1.2% of all antidepressant prescriptions), the prescription was classified as reference positive for each of the indications.

6.5.4.1.2 Quebec health administrative databases

Antidepressant prescriptions were classified as positive for a given indication according to administrative data (hereafter termed 'test positive') if the patient had an ICD-9 code for the indication recorded in either physician billing (RAMQ) or hospital discharge summary (MED-ECHO) data within -/+3 days of the prescription date. ICD-9 codes for each indication (see Appendix A) were identified from code sets used in previous studies (31,127–129). For pain, codes for osteoarthritis (130) and rheumatoid arthritis (131) were also included because pain is the primary complaint among patients with these conditions (132,133). ICD-10 codes recorded in MED-ECHO from April 2006 onwards were translated to their ICD-9 equivalent using conversion tables (134). In cases where prescriptions had a diagnostic code recorded for more than one treatment indication (only 0.6% of all antidepressant prescriptions), the prescription was classified as test positive for each of the indications.

6.5.4.2 Patient and physician characteristics

Patient age and sex were determined using beneficiary information from RAMQ. Patients' level of chronic comorbidity was measured by counting the number of distinct Charlson conditions for which the patient had a corresponding diagnostic code (127) recorded in physician billing or hospital discharge summary data over the past 365 days. Prescriber characteristics were determined using information from RAMQ including physician sex, place of medical training, and number of years in practice.

6.5.5 Statistical analysis

For each indication, a separate validation study was conducted to calculate six measures of accuracy for the indication: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR-) (Table 6-1). Sensitivity and specificity indicate the ability of a test to correctly identify those with and without the disease, while the PPV, NPV, LR+, and LR- convey information about the predictive properties of a test. Of all the accuracy measures, only the PPV and NPV are directly influenced by the disease prevalence in the study population (135–137). In general, as disease prevalence

increases, PPV increases (improves) while NPV decreases (worsens). Likelihood ratios, on the other hand, measure the predictive utility of a test without being influenced by disease prevalence.

A two-stage cluster bootstrap (103) was used to correct the 95% confidence intervals (CIs) around all accuracy measures for multi-level clustering of prescriptions within patients, who in turn were nested within physicians.

6.5.5.1 Subgroup analyses

For treatment indications with an overall prevalence of >1% according to the reference standard, subgroup analyses were conducted by antidepressant class (selective serotonin re-uptake inhibitor [SSRI], serotonin-norepinephrine reuptake inhibitor [SNRI], tricyclic antidepressant [TCA], trazodone, bupropion, or mirtazapine), patient age (<65 versus 65+ years), level of chronic comorbidity (0 versus 1+ Charlson condition), and therapy status (new versus ongoing therapy with antidepressants). Prescriptions for new antidepressant therapy were defined as prescriptions where the patient had not been prescribed an antidepressant in the MOXXI system over the past 365 days.

6.5.5.2 Sensitivity analyses

Sensitivity analyses were conducted to investigate the effect of (a) increasing the length of the lookback window to gather diagnostic codes recorded over the past 30, 60, 90, 180, or 365 days, and (b) restricting the source of diagnostic codes to hospital discharge summary data only, billings data only, or billings from the prescriber only (within a lookback window of 365 days).

To investigate how much of the total variation was due to differences in coding practices between physicians, the 95% CIs corrected for both within-patient and within-physician clustering were compared to 95% CIs corrected for within-patient clustering only. All analyses were conducted using SAS software, version 9.4.

6.6 Results

A total of 77,700 antidepressant prescriptions were written by 164 physicians for 17,606 patients. There were equal numbers of male (n=82, 50.0%) and female (n=82, 50%) prescribers, most physicians (n=150, 91.5%) had received their medical training in Canada or the US, and 76.6% of physicians (n=126) had been practicing for at least 15 years. Two-thirds of patients were female (n=11 892, 67.7%) and over the study period, each patient had a median of 3 (IQR 1-6) antidepressant prescriptions. At the time of their earliest antidepressant prescription, most patients were middle-aged (median of 53 years, IQR 43-65) and nearly one-third (n=5 404, 30.7%) had at least one condition included in the Charlson comorbidity index. Among all antidepressant prescriptions, 39.4% (n=30 596) were initiating a new therapy with antidepressants. The most commonly prescribed drugs were SSRIs (n=33 139, 42.7%), followed by SNRIs (n=18 271, 23.5%), TCAs (n=8 501, 10.9%), trazodone (n=7 216, 9.3%), bupropion (n=5 989, 7.7%), and mirtazapine (n=4 437, 5.7%). Only a very small proportion of prescriptions (<0.2%) were written for monoamine oxidase inhibitors (n=119), maprotiline (n=18), or nefazodone (n=10).

According to the indications recorded in the MOXXI e-prescribing system (reference standard), antidepressants were most commonly prescribed for depressive disorders (56.3%), anxiety/stress disorders (22.8%), sleeping disorders (10.0%), and pain (5.7%) (Table 6-2). In comparison, the proportion of antidepressant prescriptions that had diagnostic codes recorded for these indications ('test positive') was considerably lower, especially for depressive and sleeping disorders (Table 6-2). Consequently, the sensitivity of administrative diagnostic codes was very poor for all treatment indications, ranging from a high of only 31.2% (95% CI, 26.8% - 35.9%) for anxiety/stress disorders to as low as 1.3% (95% CI, 0.0% - 5.2%) for sexual dysfunction (Table 6-3). However, the specificity of diagnostic codes was excellent (90% or above) for all treatment indications (Table 6-3).

The predictive value of having a diagnostic code for a given indication recorded in administrative data varied between treatment indications. When a diagnostic code for a given indication was recorded, the probability that the antidepressant was truly prescribed for the corresponding indication (i.e. according to the MOXXI system) was high for depressive disorders (PPV of 80.3%; 95% CI, 73.7% - 85.3%), moderate for obsessive-compulsive disorder (OCD) (69.1%; 95% CI, 51.7% - 83.3%) and low (~50% or less) for the remaining indications (Table 6-3). The high PPV of depression codes was largely attributable to the high prevalence of this indication, resulting in a baseline probability of 56.3% for depressive disorders even before considering diagnostic codes. For OCD codes, however, the moderate PPV of 69.1% was not attributable to a high baseline probability, as the prevalence of OCD was only 1.1%. This contrast between codes for depressive disorders and OCD was displayed by the LR+ (ratio of true positive rate to false positive rate) because this measure is not influenced by disease prevalence. For depressive disorders, the LR+ of diagnostic codes was only 3.2 (95% CI, 2.3 – 4.4), suggesting that these codes were not very informative for ruling in the indication (Table 6-3). For OCD, however, the LR+ of diagnostic codes was 203.8 (95% CI, 103.1 – 452.2), suggesting that these codes were very informative for ruling in the indication. The extremely high LR+ for OCD codes was due to a very low false positive rate (0.1%) rather than a high true positive rate (14.9%). Codes for other indications including attention deficit/hyperactivity disorder, fibromyalgia, migraine, and nicotine dependence also had a low false positive rate, resulting in a high LR+ (Table 3). For sexual dysfunction, pre-menstrual dysphoric disorder, and eating disorders, the LR+ estimates were unstable (as reflected by the very wide 95% CIs) due to a low number of true positive and false positive prescriptions for these indications (Table 6-2).

As with the PPV and LR+, conclusions about the predictive value of *not* having a diagnostic code for a given indication differed depending on whether the NPV or LR- was used as the performance statistic. When a diagnostic code for a given indication was not recorded, the probability that the antidepressant was not prescribed for the corresponding indication according to the MOXXI system was low for depressive disorders (NPV of 49.2%; 95% CI, 45.3% - 53.2%) but fairly high for anxiety/stress disorders (81.6%; 95% CI, 78.8% - 84.0%) and high for

sleeping disorders (90.4%; 95% CI, 88.2% - 92.4%). For the remaining indications, the NPV was very high (>95%) due to the low prevalence of these indications (Table 6-3). In contrast, all estimates of the LR- (ratio of false negative rate to true negative rate) were close to 1.0 suggesting that for all indications, the absence of a diagnostic code for a given indication did not improve the ability to rule out the corresponding indication.

6.6.1 Subgroup analyses

For all indications, there was considerable heterogeneity in the PPV and NPV estimates across different classes of antidepressants (Table 6-4). Diagnostic codes usually had better PPV and poorer NPV for antidepressants with a higher prevalence or baseline probability of the indication. For example, mirtazapine was far more likely to be prescribed for depression than trazodone (baseline probability of 86.9% versus 10.4%), thus explaining why diagnostic codes for depression had a much higher PPV for mirtazapine (96.5% versus 20.3%) but a much lower NPV (15.1% versus 91.3%). However, there were two exceptions to this trend. For fibromyalgia, the baseline probability of this indication was similar for SNRIs and TCAs (3.1% versus 3.4%) but the PPV for SNRIs (62.7%; 95% CI, 47.5% - 76.0%) was much higher than for TCAs (32.0%; 95% CI, 16.9% - 45.9%). Similarly, the baseline probability of OCD was low for both SSRIs and SNRIs (2.0% versus 0.7%), yet the PPV for SSRIs (81.0%; 95% CI, 62.5% - 94.4%) was much higher than for SNRIs (38.1%; 95% CI, 0.0% - 77.8%). Unlike the PPV and NPV, the LR+ and LR- estimates did not differ notably across antidepressant classes (data not shown), thus demonstrating that these measures are not influenced by disease prevalence.

When prescriptions were stratified by patients' level of chronic comorbidity and age, diagnostic codes for all indications had noticeably poorer sensitivity (and consequently, slightly better specificity) among patients with at least one comorbidity in the Charlson index and patients 65+ years old, especially for depressive disorders and anxiety/stress disorders (Tables 6-5 and 6-6). The stratum-specific estimates for sicker and older patients were similar because these two characteristics were strongly correlated (data not shown). Among prescriptions for new versus ongoing antidepressant therapy, the sensitivity and PPV of diagnostic codes was better among

prescriptions for new antidepressant therapy for all indications except depressive disorders and fibromyalgia (Table 6-7).

6.6.2 Sensitivity analyses

As one would expect, using a longer lookback window for diagnostic codes increased sensitivity and decreased specificity for all indications, especially pain (Figure 6-1, panels A and B). However, even with the longest lookback window of 365 days, sensitivity remained low at ≤60% for all indications. Increasing the length of the lookback window caused the PPV to deteriorate for all indications, especially sleeping disorders (Figure 6-1, panel C), while NPV remained quite stable (Figure 6-1, panel D). As a result of these trends, the LR+ decreased (worsened) with a longer lookback window for all indications, especially sleeping disorders, migraine, fibromyalgia, and OCD (Figure 6-1, Panel E), while the LR- decreased (improved) for all indications, especially pain (Figure 6-1, Panel F).

Compared to the performance of diagnostic codes from physician billings when using a lookback window of 365 days, diagnostic codes from hospital discharge data in the past 365 days produced drastically lower sensitivities (and thus higher specificities) for all indications (Figure 6-2, panels A and B). Hospital diagnostic codes also had worse predictive performance than diagnostic codes from billings data, as shown by the lower PPV, NPV and LR+ values, and higher LR- values (Figure 6-2, panels C-F).

When diagnostic codes from billings in the past 365 days were restricted to those from the prescribing physician only, the sensitivity of diagnostic codes for most indications was only slightly lower (and consequently, specificity was slightly higher) relative to using billings from all physicians (Figure 6-2, panels A and B). For pain, however, the sensitivity of diagnostic codes from the prescribing physician versus all physicians was drastically lower (42% versus 60%). Diagnostic codes recorded by the prescribing physician had slightly better predictive performance (higher PPV and LR+) than diagnostic codes recorded by all physicians (Figure 6-2, panels C and E). Diagnostic codes from the prescribing physician also produced LR-s that were

slightly higher (worse) than diagnostic codes from all physicians (Figure 6-2, panel F), indicating that the absence of a code for a given indication among billings from all physicians was more informative for ruling out the indication than not having a code for the indication recorded among billings from the single physician who prescribed the drug.

Finally, for all indications except sleeping disorders, the 95% cluster bootstrap-based CIs (103) around the sensitivity and PPV estimates were noticeably wider when they accounted for both within-physician and within-patient clustering than when they accounted for within-patient clustering only (Figure 6-3). In fact, for depressive and anxiety/stress disorders, the patient-only-adjusted 95% CIs were much narrower, indicating that failure to correct for between-physician differences in coding accuracy would have considerably underestimated the variance around the point estimate.

6.7 Discussion

In this study, we estimated the accuracy with which diagnostic codes in Quebec health administrative records reflected indications for antidepressant therapy in primary care. We found that diagnostic codes for a given indication identified only a small proportion of antidepressant prescriptions for the indication. Moreover, we found that the absence of a diagnostic code for a given indication did not provide much additional value for ruling out the indication.

The findings from this validation study have important implications for epidemiological studies using administrative diagnostic codes as proxies for antidepressant treatment indications. Studies aimed at monitoring rates of antidepressant use for off-label indications will significantly over-estimate the true off-label prescribing rate since a large proportion of truly on-label antidepressant prescriptions will not have a diagnostic code recorded for the corresponding indication. In fact, a study of Georgia Medicaid enrollees (34) concluded that 75% of antidepressant recipients in 2001 received an antidepressant off-label because a diagnostic code for an approved indication could not be found for the patient within the same

year. In comparison, using our dataset of physician-documented treatment indications, we previously found that the proportion of antidepressants prescribed for off-label indications was considerably lower, at only 29% (91,121). Our findings also suggest that in studies evaluating the safety of off-label antidepressant use, the use of administrative diagnostic codes to infer treatment indications could misclassify a significant proportion of on-label users as off-label users, thus possibly diluting or even concealing adverse drug events among off-label users. For example, in the case of trazodone (approved in Canada for depression only), we found that diagnostic codes for depressive disorders had an NPV of 91.3% (95% CI, 85.9% - 95.5%), suggesting that 8.7% (95% CI, 4.5% - 14.1%) of supposedly off-label users could in fact be on-label users (since they do not have a diagnostic code for depression but have been prescribed trazodone to treat depression). For mirtazapine (also approved in Canada for depression only), the NPV of diagnostic codes for depression was much lower at 15.1% (95% CI, 10.2% - 21.0%), suggesting that 84.9% (95% CI, 79.0% - 89.8%) of off-label users could be misclassified. These examples are just a few ways that the accuracy estimates from this study could be used to inform bias analyses in epidemiological studies where antidepressant treatment indications are inferred using administrative diagnostic codes.

Our study demonstrates the strong influence that disease prevalence can have on the PPV and NPV. Among antidepressants with a higher baseline prevalence of a given indication, diagnostic codes for the indication generally had better PPV and worse NPV than among antidepressants with a lower baseline prevalence of the indication. Thus, when reporting results in validation studies, our findings demonstrate the importance of not only describing the disease prevalence in the study population, but also stratifying the analysis by factors that are expected to affect disease prevalence in the study population. Furthermore, unlike the PPV and NPV, the fact that the LR+ and LR- estimates did not differ notably across antidepressant classes demonstrates that these measures are not influenced by disease prevalence. As such, it may be useful to report likelihood ratios in addition to PPV and NPV estimates when assessing the predictive properties of a diagnostic test. Likelihood ratios have traditionally received little attention in validation studies due to their seemingly complex interpretation involving nomograms and

converting between the probability and odds of a disease (136,138). However, likelihood ratios can be thought of as simply the probability of a particular test result among patients with versus without the disease of interest (139). Using this definition, likelihood ratios become useful measures for quickly assessing the added value of a diagnostic test for ruling in or ruling out disease, regardless of its prevalence.

Several reasons may help explain the poor accuracy of diagnostic codes from Quebec health administrative data. First, physicians have little incentive to accurately record diagnostic codes when completing medical claims since physicians are not even required to submit a diagnostic code. Second, since only one diagnosis can be recorded per claim, this limitation reduces the likelihood that a code for the antidepressant treatment indication will be captured, especially among patients with multiple morbidities. Indeed, we found that administrative diagnostic codes had lower sensitivity among patients with higher levels of chronic comorbidity. Third, depressed patients may often be hospitalized for other medical conditions, leaving their depression overlooked, under recognized, and rarely captured in the hospital discharge diagnoses (140). Given the poor accuracy of administrative diagnostic codes, future work should determine whether the ability of administrative data to predict antidepressant treatment indications can be improved by considering other patient, physician, and drug-related factors in addition to diagnostic codes.

Our finding that the sensitivity of diagnostic codes for pain was much lower when restricted to billings from the prescribing physician compared to billings from all physicians suggests that patients who are prescribed antidepressants for pain are likely to see multiple physicians for their pain. For all other indications, however, the fact that the sensitivity of diagnostic codes from the prescribing physician was only slightly lower than using codes from all physicians suggests that when primary care physicians prescribe antidepressants for indications other than pain, they may often provide the majority of care for the corresponding indication.

Our finding that diagnostic codes for most indications had slightly better accuracy among prescriptions for new versus ongoing antidepressant therapy was not surprising. When antidepressants are first prescribed, the indication for the new prescription may be among the primary complaints responsible for the visit, thus increasing the probability that a diagnostic code for the indication is recorded. For depressive disorders and fibromyalgia, however, our finding that diagnostic codes for these indications did not have better accuracy among new prescriptions was unexpected but consistent with reports that primary care physicians often have difficulty recognizing and diagnosing these conditions in the initial stages (141,142). As a result, physicians may be less likely to record diagnostic codes for these indications when antidepressants are first prescribed.

This study has potential limitations that should be considered. First, the treatment indications we validated in this study accounted for the indications of 99.5% of antidepressant prescriptions in the MOXXI system. For the remaining 0.5% of prescriptions, we did not validate the indications (e.g. fatigue, bipolar disorder, obesity, Crohn's disease, and irritable bowel syndrome) because they were so rare. Second, the external generalizability of our findings depends on the extent to which diagnostic coding practices are similar between MOXXI physicians and physicians in other settings. MOXXI physicians operate within a universal, publicly-funded health care system where all citizens are insured for essential medical care regardless of health or mental status. In the US, however, where health care delivery is heavily privatized, physicians may be more cautious about recording diagnostic codes for psychiatric conditions. In fact, primary care physicians in the US have reported commonly recording alternative diagnoses such as fatigue, insomnia, or headache for patients with major depression due to concerns over obtaining reimbursement or jeopardizing patients' future ability to obtain health insurance (104). Consequently, the accuracy of administrative diagnostic codes for antidepressant treatment indications in the US may be even worse than in Canada. Another limitation of our study is that we could not determine how often MOXXI physicians recorded only one indication in the e-prescribing system when there were truly multiple indications responsible for the prescription. If certain indications were often omitted, then we may have

overestimated the NPV and underestimated the PPV of diagnostic codes for these indications. We hypothesize that this phenomenon did not occur too frequently, as the physicians in our study documented multiple indications for 1.8% of antidepressant prescriptions, which is only slightly lower than the 5% of adult antidepressant users in a UK study who reported taking antidepressants for multiple indications (16). Finally, in our main analysis, we used a short lookback window of three days for diagnostic codes because we knew when the index prescription was written. However, in cases where only dispensing data is available, a longer lookback window may be necessary in order to capture the index visit, especially since some antidepressants (e.g. SSRIs) may be taken on an 'as-needed' basis for certain indications (15).

In conclusion, the findings from this study suggest that diagnostic codes from administrative data should not be used alone to infer antidepressant treatment indications. Future studies should determine whether diagnostic codes can be combined with other information from administrative databases to improve the ability to accurately predict antidepressant treatment indications.

Table 6-1. Measures of accuracy for each antidepressant treatment indication

		Reference standard (MOXXI)	
		Positive for the indication	**Negative for the indication**
Administrative data	**Positive for the indication**	True positive (TP)	False positive (FP)
	Negative for the indication	False negative (FN)	True negative (TN)

Measure	**Formula**	**Interpretation using depression as an example**
Sensitivity[a]	TP/(TP+FN)	Probability that an antidepressant prescription for depression has a diagnostic code for depression recorded.
Specificity[a]	TN/(TN+FP)	Probability that an antidepressant prescription not for depression does not have a diagnostic code for depression recorded.
Positive predictive value[a] (PPV)	TP/(TP+FP)	Probability that an antidepressant prescription with a code for depression is truly for depression.
Negative predictive value[a] (NPV)	TN/(TN+FN)	Probability that an antidepressant prescription without a code for depression is truly not for depression.
Positive likelihood ratio[a] (LR+)	Sensitivity/(1-Specificity)	The number of times more likely it is that a diagnostic code for depression is recorded among prescriptions for depression compared to prescriptions not for depression. Tests with a LR+ of 10 or greater are often considered as having high diagnostic value (139).
Negative likelihood ratio[b] (LR-)	(1-Sensitivity)/Specificity	The number of times more likely it is that a diagnostic code for depression is not recorded among prescriptions for depression compared to prescriptions not for depression. Tests with a LR- of 0.1 or less are often considered as having high diagnostic value (139).

[a]Higher values indicate better performance of diagnostic codes for a given indication.
[b]Lower values indicate better performance of diagnostic codes for a given indication.

Table 6-2. Proportion of antidepressant prescriptions for each treatment indication according to MOXXI and Quebec health administrative data

	Number (%) of antidepressant prescriptions							
Treatment indication	MOXXI[a] [reference standard]		Quebec health administrative data[b]		TP	TN	FN	FP
Depressive disorders	43,752	(56.3)	14,465	(18.6)	11,610	31,093	32,142	2,855
Anxiety/stress disorders	17,677	(22.8)	11,606	(14.9)	5,520	53,937	12,157	6,086
Sleeping disorders	7,771	(10.0)	720	(0.9)	380	69,589	7,391	340
Pain	4,416	(5.7)	4,090	(5.3)	847	70,041	3,569	3,243
Migraine	1,162	(1.5)	737	(1.0)	259	76,060	903	478
Fibromyalgia	917	(1.2)	796	(1.0)	256	76,243	661	540
Obsessive-compulsive disorder	840	(1.1)	181	(0.2)	125	76,804	715	56
Vasomotor symptoms of menopause	599	(0.8)	613	(0.8)	48	76,536	551	565
Nicotine dependence	432	(0.6)	108	(0.1)	18	77,178	414	90
Attention deficit/hyperactivity disorder	255	(0.3)	119	(0.2)	23	77,349	232	96
Sexual dysfunction	228	(0.3)	10	(0.0)	3	77,465	225	7
Pre-menstrual dysphoric disorder	146	(0.2)	26	(0.0)	9	77,537	137	17
Eating disorders	74	(0.1)	31	(0.0)	9	77,604	65	22

Abbreviations: MOXXI = Medical Office of the XXIst Century, TP = True Positive, TN = True Negative, FN = False Negative, FP = False Positive

[a]Based on physician-documented treatment indications recorded for antidepressant prescriptions in the MOXXI system. 1.2% of prescriptions were classified as reference positive for multiple treatment indications because more than one indication was recorded for the prescription in the MOXXI system.

[b]Based on diagnostic codes in physician billing and hospitalization discharge summary data that were recorded for patients within -/+ 3 days of the prescription date. 0.6% of prescriptions were classified as test positive for multiple treatment indication because diagnostic codes for more than one treatment indication were recorded.

Table 6-3. Accuracy of diagnostic codes from Quebec health administrative databases for identifying antidepressant treatment indications

Treatment indication	Prevalence, %	Sensitivity, % (95% CI)		Specificity, % (95% CI)		PPV, % (95% CI)		NPV, % (95% CI)		LR+ (95% CI)		LR- (95% CI)	
Depressive disorders	56.3	26.5	(20.7-32.0)	91.6	(87.6-94.6)	80.3	(73.7-85.3)	49.2	(45.3-53.2)	3.2	(2.3-4.4)	0.80	(0.75-0.85)
Anxiety/stress disorders	22.8	31.2	(26.8-35.9)	89.9	(87.1-92.3)	47.6	(41.8-54.3)	81.6	(78.8-84.0)	3.1	(2.5-3.9)	0.77	(0.72-0.81)
Sleeping disorders	10.0	4.9	(3.4-6.8)	99.5	(99.3-99.7)	52.8	(46.0-60.1)	90.4	(88.2-92.4)	10.1	(7.2-14.7)	0.96	(0.94-0.97)
Pain	5.7	19.2	(15.5-23.0)	95.6	(94.8-96.3)	20.7	(16.4-25.9)	95.2	(94.2-95.9)	4.3	(3.5-5.4)	0.85	(0.81-0.88)
Migraine	1.5	22.3	(17.0-29.1)	99.4	(99.2-99.5)	35.1	(26.2-45.2)	98.8	(98.4-99.2)	35.7	(27.0-49.2)	0.78	(0.71-0.84)
Fibromyalgia	1.2	27.9	(18.8-38.8)	99.3	(99.0-99.5)	32.2	(23.9-40.2)	99.1	(98.9-99.4)	39.7	(28.1-55.5)	0.73	(0.62-0.82)
Obsessive-compulsive disorder	1.1	14.9	(7.5-23.4)	99.9	(99.9-100.0)	69.1	(51.7-83.3)	99.1	(98.8-99.3)	203.8	(103.1-452.2)	0.85	(0.77-0.93)
Vasomotor symptoms of menopause	0.8	8.0	(3.8-13.3)	99.3	(98.9-99.5)	7.8	(3.5-14.3)	99.3	(99.0-99.5)	10.9	(5.1-20.7)	0.93	(0.87-0.97)
Nicotine dependence	0.6	4.2	(0.7-9.4)	99.9	(99.8-99.9)	16.7	(4.2-29.3)	99.5	(99.2-99.7)	35.9	(8.2-73.8)	0.96	(0.91-0.99)
Attention deficit/hyperactivity disorder	0.3	9.0	(2.1-17.3)	99.9	(99.8-99.9)	19.3	(5.3-37.1)	99.7	(99.5-99.8)	72.7	(20.3-178.8)	0.91	(0.83-0.98)
Sexual dysfunction	0.3	1.3	(0.0-5.2)	100.0	(100.0-100.0)	30.0	(0.0-88.9)	99.7	(99.5-99.9)	146.2	(0.0-1337.5)	0.99	(0.95-1.00)
Pre-menstrual dysphoric disorder	0.2	6.2	(0.0-15.4)	100.0	(99.9-100.0)	34.6	(0.0-71.4)	99.8	(99.7-99.9)	280.2	(0.0-1434.5)	0.94	(0.84-1.00)
Eating disorders	0.1	12.2	(0.0-32.8)	100.0	(99.9-100.0)	29.0	(0.0-66.7)	99.9	(99.8-100.0)	434.4	(0.0-2111.6)	0.88	(0.67-1.00)

Abbreviations: PPV = positive predictive value, NPV = negative predictive value, LR+ = positive likelihood ratio, LR- = negative likelihood ratio

Table 6-4. Positive predictive value (PPV) and negative predictive value (NPV) of administrative diagnostic codes for the seven most common treatment indications, by antidepressant class

Treatment indication, by antidepressant class[a]	Prevalence, %	PPV, % (95% CI)		NPV, % (95% CI)	
Depressive disorders					
SSRI	61.9	86.1	(82.2-89.4)	43.4	(38.6-48.7)
SNRI	67.1	88.3	(84.5-91.4)	39.3	(34.0-44.5)
TCA	14.7	27.3	(13.2-51.9)	86.4	(82.6-89.8)
Trazodone	10.4	20.3	(9.6-36.0)	91.3	(85.9-95.5)
Bupropion	84.1	93.1	(89.0-97.0)	19.5	(14.0-25.1)
Mirtazapine	86.9	96.5	(93.5-98.6)	15.1	(10.2-21.0)
Anxiety/stress disorders					
SSRI	36.0	63.2	(56.0-70.4)	70.2	(65.3-74.4)
SNRI	24.1	51.3	(41.9-62.0)	80.8	(77.3-84.0)
TCA	3.2	7.1	(2.9-14.0)	97.1	(95.6-98.3)
Trazodone	7.8	8.9	(4.4-14.3)	92.4	(89.0-95.0)
Bupropion	0.3	1.2	(0.0-3.2)	99.8	(99.6-100.0)
Mirtazapine	10.3	15.3	(9.0-26.6)	90.5	(84.8-94.8)
Sleeping disorders					
SSRI	0.0	0.0	(0.0-0.0)	100.0	(99.9-100.0)
SNRI	0.0	1.5	(0.0-6.8)	100.0	(100.0-100.0)
TCA	20.0	67.5	(44.7-84.3)	80.5	(72.2-87.5)
Trazodone	82.0	95.8	(90.8-98.8)	18.6	(13.1-25.4)
Bupropion	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)
Mirtazapine	3.2	21.0	(3.5-31.5)	97.4	(95.0-99.3)
Pain					
SSRI	0.1	0.4	(0.0-1.1)	100.0	(99.9-100.0)
SNRI	3.1	15.7	(9.3-23.4)	97.6	(96.7-98.3)
TCA	42.8	72.0	(62.9-79.6)	60.8	(53.2-67.5)
Trazodone	1.6	4.4	(0.9-9.6)	98.6	(97.4-99.4)
Bupropion	1.3	5.1	(0.3-11.9)	98.8	(97.9-99.5)
Mirtazapine	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)

Abbreviations: SSRI = selective serotonin reuptake inhibitor; SNRI = serotonin-norepinephrine reuptake inhibitor; TCA = tricyclic antidepressant

[a]SSRIs include citalopram, paroxetine, sertraline, escitalopram, fluoxetine, and fluvoxamine. SNRIs include venlafaxine, duloxetine, and desvenlafaxine. TCAs include amitriptyline, doxepin, trimipramine, nortriptyline, imipramine, clomipramine, and desipramine. Results are not shown for monoamine oxidase inhibitors, maprotiline, or nefazodone due to small numbers of prescriptions for each of these drugs.

Table 6-4. (continued) Positive predictive value (PPV) and negative predictive value (NPV) of administrative diagnostic codes for the seven most common treatment indications, by antidepressant class

Treatment indication, by antidepressant class[a]	Prevalence, %	PPV, % (95% CI[b])		NPV, % (95% CI[b])	
Migraine					
SSRI	0.0	2.2	(0.0-7.2)	100.0	(99.9-100.0)
SNRI	0.0	0.0	(0.0-0.0)	100.0	(99.9-100.0)
TCA	13.5	71.6	(61.5-80.5)	89.0	(85.1-92.4)
Trazodone	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)
Bupropion	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)
Mirtazapine	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)
Fibromyalgia					
SSRI	0.1	7.9	(0.0-20.5)	99.9	(99.7-100.0)
SNRI	3.1	62.7	(47.5-76.0)	97.7	(96.8-98.5)
TCA	3.4	32.0	(16.9-45.9)	97.5	(96.3-98.6)
Trazodone	0.0	1.4	(0.0-7.3)	100.0	(99.9-100.0)
Bupropion	0.3	30.0	(0.0-64.3)	99.9	(99.6-100.0)
Mirtazapine	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)
Obsessive-compulsive disorder					
SSRI	2.0	81.0	(62.5-94.4)	98.3	(97.8-98.8)
SNRI	0.7	38.1	(0.0-77.8)	99.3	(98.7-99.7)
TCA	0.4	85.7	(0.0-100.0)	99.6	(99.2-99.9)
Trazodone	0.0	0.0	(0.0-0.0)	100.0	(100.0-100.0)
Bupropion	0.1	0.0	(0.0-0.0)	100.0	(99.8-100.0)
Mirtazapine	0.0	N/A[b]		100.0	(100.0-100.0)

Abbreviations: SSRI = selective serotonin reuptake inhibitor; SNRI = serotonin-norepinephrine reuptake inhibitor; TCA = tricyclic antidepressant

[a]SSRIs include citalopram, paroxetine, sertraline, escitalopram, fluoxetine, and fluvoxamine. SNRIs include venlafaxine, duloxetine, and desvenlafaxine. TCAs include amitriptyline, doxepin, trimipramine, nortriptyline, imipramine, clomipramine, and desipramine. Results are not shown for monoamine oxidase inhibitors, maprotiline, or nefazodone due to small numbers of prescriptions for each of these drugs.

[b]Could not be calculated due to a zero denominator (i.e. no prescriptions for mirtazapine had a diagnostic code for obsessive-compulsive disorder recorded within -/+ 3 days of the prescription date).

Table 6-5. Sensitivity and specificity of diagnostic codes for the seven most common treatment indications, by level of patient chronic comorbidity

Treatment indication	Sensitivity, % (95% CI)				Specificity, % (95% CI)			
	0 Charlson conditions		1+ Charlson conditions		0 Charlson conditions		1+ Charlson conditions	
Depressive disorders	31.2	(24.3-37.8)	16.7	(12.8-20.3)	90.2	(85.2-93.9)	94.7	(92.5-96.6)
Anxiety/stress disorders	35.8	(30.4-41.0)	19.5	(16.0-23.5)	88.3	(85.3-91.3)	93.0	(91.2-94.6)
Sleeping disorders	5.0	(3.3-7.1)	4.7	(2.9-6.4)	99.5	(99.3-99.7)	99.5	(99.2-99.7)
Pain	21.0	(16.6-25.4)	16.3	(11.3-21.0)	95.9	(95.2-96.6)	94.8	(93.8-95.8)
Migraine	25.3	(18.9-33.3)	12.6	(7.0-21.2)	99.3	(99.0-99.5)	99.6	(99.4-99.7)
Fibromyalgia	33.0	(21.9-43.2)	16.9	(6.6-29.5)	99.2	(98.9-99.5)	99.4	(99.2-99.7)
Obsessive-compulsive disorder	16.4	(8.5-26.2)	8.5	(0.5-21.5)	99.9	(99.8-100.0)	100.0	(99.9-100.0)

Table 6-6. Sensitivity and specificity of diagnostic codes for the seven most common treatment indications, by patient age

Treatment indication	Sensitivity, % (95% CI)				Specificity, % (95% CI)			
	<65 years		65+ years		<65 years		65+ years	
Depressive disorders	29.9	(23.6-36.5)	16.0	(12.8-20.0)	89.9	(85.3-93.5)	96.1	(94.1-97.5)
Anxiety/stress disorders	35.7	(30.7-40.4)	18.4	(14.2-23.3)	88.6	(85.6-91.4)	93.7	(92.0-95.3)
Sleeping disorders	5.1	(3.3-7.3)	4.5	(2.9-6.5)	99.5	(99.2-99.7)	99.6	(99.4-99.8)
Pain	20.7	(16.4-25.0)	16.0	(11.0-21.5)	95.9	(95.2-96.7)	94.4	(93.5-95.4)
Migraine	24.1	(17.6-31.7)	13.0	(4.8-24.0)	99.3	(99.0-99.5)	99.7	(99.5-99.8)
Fibromyalgia	29.6	(18.8-40.5)	20.2	(6.1-36.0)	99.2	(98.9-99.5)	99.6	(99.4-99.8)
Obsessive-compulsive disorder	17.0	(9.0-27.4)	2.4	(0.0-9.9)	99.9	(99.8-100.0)	100.0	(99.9-100.0)

Table 6-7. Sensitivity and positive predictive value (PPV) of diagnostic codes for the seven most common treatment indications, by antidepressant therapy status

Treatment indication	Sensitivity, % (95% CI)				PPV, % (95% CI)			
	New therapy[a]		Ongoing therapy		New therapy[a]		Ongoing therapy	
Depressive disorders	26.1	(21.7-29.9)	26.8	(20.2-33.3)	80.8	(72.6-87.0)	80.0	(74.0-84.8)
Anxiety/stress disorders	33.7	(28.5-38.3)	29.5	(24.8-34.3)	52.4	(45.5-59.7)	44.3	(38.0-52.0)
Sleeping disorders	6.6	(4.5-9.0)	3.6	(2.1-5.3)	61.9	(52.7-71.5)	43.8	(34.4-53.5)
Pain	20.4	(16.2-24.3)	18.2	(13.4-23.1)	26.1	(20.6-31.4)	17.4	(12.6-23.0)
Migraine	28.0	(21.5-34.7)	16.3	(10.1-25.9)	48.0	(37.0-58.8)	23.7	(15.0-34.9)
Fibromyalgia	26.3	(16.6-36.3)	28.7	(18.1-40.8)	32.4	(22.2-42.7)	32.0	(22.7-42.3)
Obsessive-compulsive disorder	17.0	(7.5-28.1)	13.4	(5.3-23.3)	71.6	(51.9-87.2)	67.0	(44.3-85.4)

[a]Defined as prescriptions where the patient had not been prescribed an antidepressant in the MOXXI system over the past 365 days.

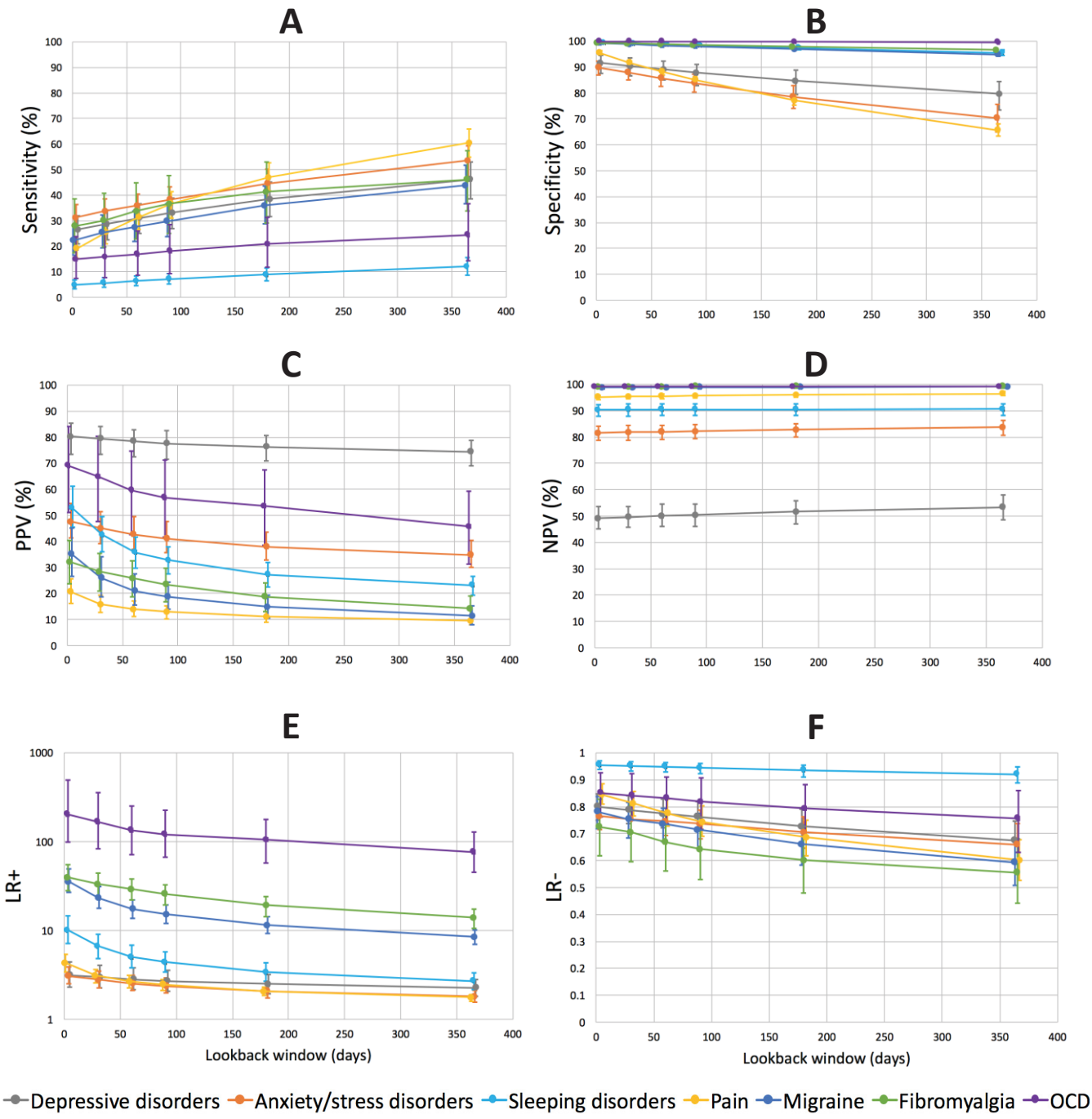


Figure 6-1. Effect of increasing the lookback window for administrative diagnostic codes

The figure shows the accuracy of administrative diagnostic codes for the seven most common treatment indications recorded in the past 3, 30, 60, 90, 180, and 365 days. Abbreviations: PPV = positive predictive value; NPV = negative predictive value; LR+ = positive predictive value; LR- = negative predictive value.

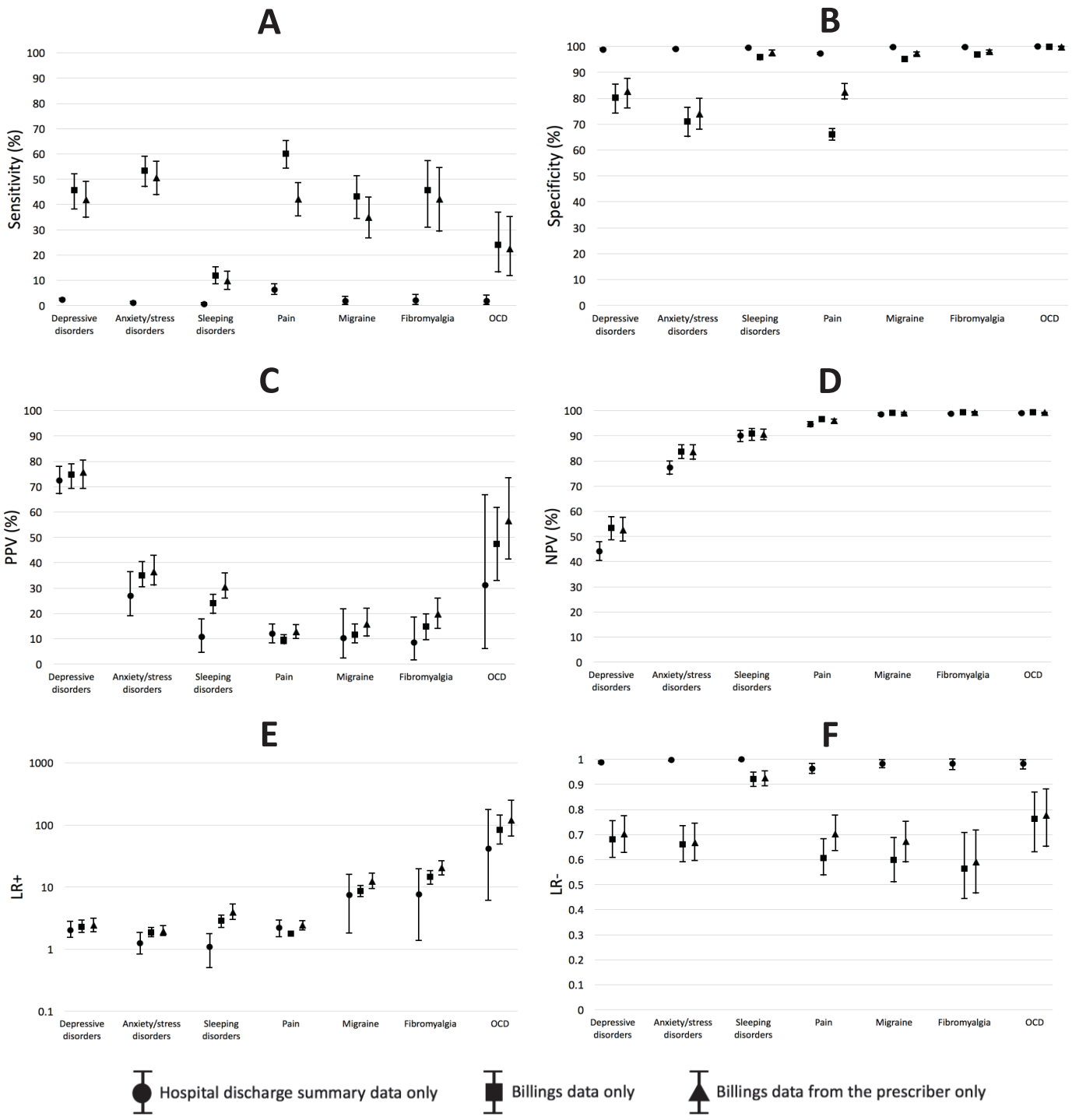


Figure 6-2. Effect of restricting diagnostic codes to different sources of administrative data

The figure shows the accuracy of diagnostic codes for the seven most common treatment indications recorded within the past 365 days when restricted to diagnostic codes from either hospital discharge summary data, billings from all physicians, or billings from the prescribing physician. Abbreviations: PPV = positive predictive value; NPV = negative predictive value; LR+ = positive predictive value; LR- = negative predictive value.

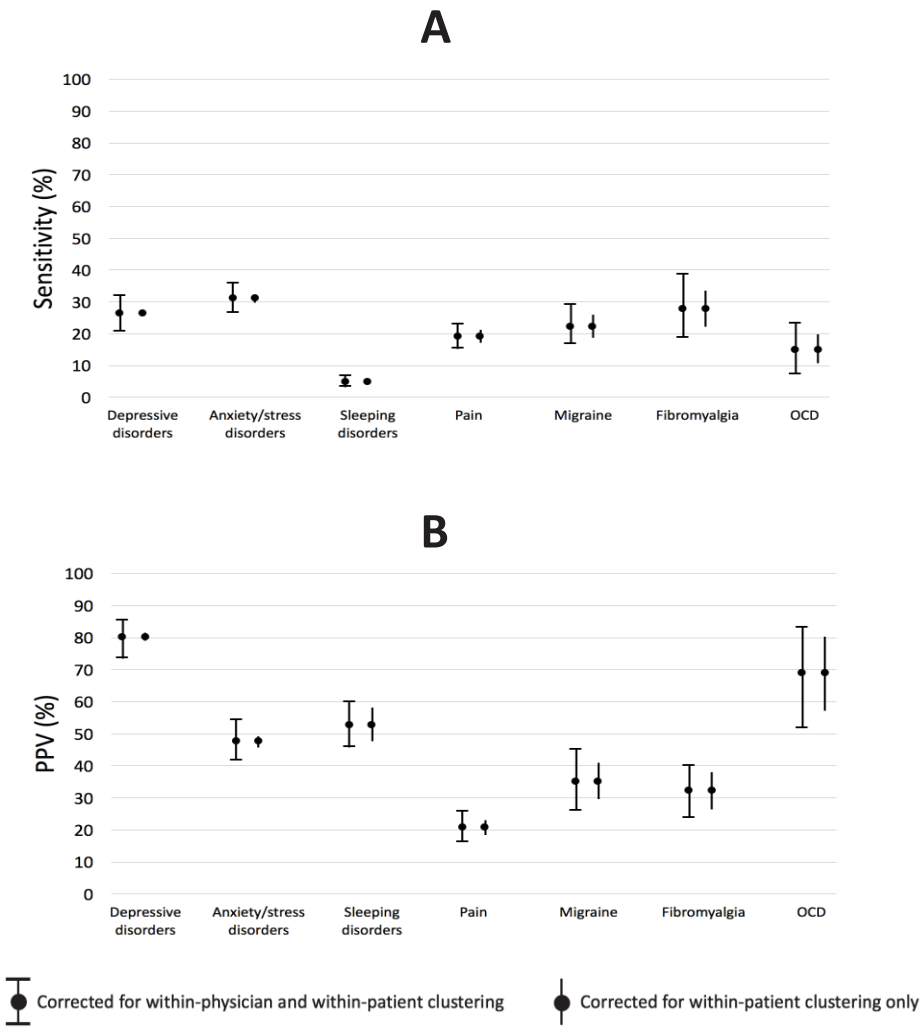


Figure 6-3. Variance of sensitivity and positive predictive value (PPV) estimates when corrected for both within-physician and within-patient clustering versus within-patient clustering only

The figure shows the 95% CIs around the sensitivity (panel A) and PPV estimates (panel B) when a two-stage cluster bootstrap (103) was used to correct for clustering of prescriptions within patients who in turn were nested within physicians (capped vertical bars) versus when a one-stage cluster bootstrap was used to correct for only clustering of prescriptions within patients (uncapped vertical bars). The upper and lower bounds of the 95% CI correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples. Results are shown for the seven most common treatment indications based on diagnostic codes recorded in administrative data within -/+ 3 days of the prescription date.

7 Derivation and validation of a model to identify when antidepressants are prescribed for indications besides depression: a prediction study

7.1 Preamble

The findings from manuscript 3 showed that the use of diagnostic codes alone is not sufficient to determine antidepressant treatment indications and could introduce substantial misclassification bias in analyses where this approach is used. Better methods for measuring antidepressant treatment indications are therefore needed. The purpose of this manuscript was to address this need by building a model using standard regression techniques to predict when antidepressants were prescribed for indications besides depression based on a wide range of variables derived from health services data beyond diagnostic codes. This manuscript also identified important predictors of antidepressant prescriptions for indications besides depression and quantified their association with the outcome.

7.2 Title page and footnotes

Title: Derivation and validation of a model to identify when antidepressants are prescribed for indications besides depression: a prediction study

Authors: Jenna Wong[1], MSc; Michal Abrahamowicz[1], PhD; David Buckeridge[1], MD, PhD; and Robyn Tamblyn[1], PhD

Affiliations:

[1]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

Corresponding author:

Jenna Wong

1140 Pine Avenue West

Montreal, Quebec, Canada

H3A 1A3

Email: jenna.wong@mail.mcgill.ca

7.3 Abstract

Objectives To develop a model that uses variables in health service data to accurately predict when antidepressants are prescribed for indications besides depression.

Design Prediction study.

Setting Primary care practices in and around two major urban centres in Quebec, Canada.

Participants Primary care physicians who were prescribed antidepressants to their patients between 1 January 2003 and 31 December 2012 using an indication-based electronic prescribing system.

Main outcome measure Whether an antidepressant prescription was written for an indication besides depression.

Results The analysis included 73 576 antidepressant prescriptions written by 141 physicians for 16 262 patients. Among these prescriptions, 32 405 (44.0%) were written for indications besides depression. The final prediction model was derived using 3-fold cross-validation and a forward stepwise selection procedure. The name of the molecule prescribed was by far the strongest predictor of whether an antidepressant was prescribed for depression. Other predictors included diagnostic codes and drugs prescribed in the past year, the patient's age and education level, the physician's workload, and the prescribed dose. In the test set, the final model had good discrimination (*c* statistic: 0.815, 95% CI 0.787 to 0.847), good calibration (ratio of observed to expected events: 0.986, 95% CI 0.842 to 1.136), and performed substantially better than a model whose covariates were based on diagnostic codes only.

Conclusions We found that variables in administrative health data could be used to accurately identify when antidepressants in primary care were being prescribed for indications besides depression. Our model represents a valuable tool that could allow researchers to predict the reasons for antidepressant prescriptions in the absence of documented treatment indications, thus addressing a major barrier that prevents administrative data from being used to study antidepressant use by indication.

7.4 Introduction

Adverse drug reactions are a major cause of mortality, morbidity, and hospitalizations in the United States (143–145) and Canada.(6) 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" (8) – is essential for increasing drug safety and reducing the frequency of adverse drug events.

An important component of pharmacovigilance is monitoring the use of medicinal products for unapproved indications.(146) Antidepressants are a group of drugs in need of increased pharmacovigilance because they are among the most commonly prescribed medications (1,90,147) and patients take them for a wide variety of indications besides depression, many of which are unapproved or "off-label" for the drug. In fact, we found that nearly half of all antidepressant prescriptions in primary care were written for indications besides depression, two-thirds of which were off-label for the drug.(91) Moreover, most off-label indications for antidepressants were not backed by a strong level of scientific evidence,(121) which raises concerns because unsupported off-label drug use increases the risk of adverse drug events and places unnecessary financial burdens on the health care system.(99) The frequency with which antidepressants are prescribed for non-evidenced-based indications emphasizes 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 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.(122) 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, previous studies (17,31,34,59) have inferred the reasons for antidepressant use based on

diagnostic codes for plausible indications recorded for patients in medical or 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 for indications like depression, anxiety/stress disorders, and sleeping disorders had a sensitivity of only 26.5%, 31.2%, and 4.9%, respectively, for identifying antidepressant prescriptions for these indications.(148)

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 that used a wide range of variables derived from health services data to predict whether antidepressants were prescribed for indications besides 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 was to identify important predictors of antidepressant prescriptions for indications besides depression and measure their association with the outcome.

7.5 Methods

7.5.1 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.(149) 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.(82) 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) that contains 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 (i.e. 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.

7.5.2 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.(36) 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%).(37)

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.(80)

7.5.3 Data sources

This study included all MOXXI prescriptions of drugs approved for depression that were written 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. This study was approved by the McGill institutional review board.

7.5.4 Exclusion criteria

Drugs were excluded if they had fewer than 120 prescriptions in the MOXXI system during 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.

7.5.5 Study outcomes

The primary outcome was a binary variable representing whether an antidepressant had been prescribed for an indication besides 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.

7.5.6 Candidate predictors

There is little empirical evidence on factors associated with treatment indications for antidepressants. Two studies have identified factors associated with antidepressant users with diagnostic codes for depression (59) or other diagnoses of interest.(31) However, predictors of diagnostic codes for plausible antidepressant treatment indications may differ from predictors

of *actual* treatment indications since these two measures are poorly correlated.(148) Thus, we considered a wide range of prescription, patient, and physician-related factors that could be important predictors of antidepressant treatment indications. Table 7-1 shows all 373 variables that were included in the analysis.

7.5.6.1 Prescription factors

Using information on the index prescription, we created variables for the name of the prescribed drug, 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 (59,91) but may be taken 'as-needed' when used for other conditions like sexual dysfunction.(15) We also hypothesized that the prescribed dose would be important to consider because drugs like trazodone, amitriptyline, and doxepin are typically taken at lower-than-usual doses when prescribed for insomnia.(150)

7.5.6.2 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 age 65 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 (i.e. hot flashes), nicotine dependence, attention deficit/hyperactivity disorder, sexual dysfunction, pre-menstrual dysphoric disorder, and eating disorders. Two separate observation windows for these 13 diagnostic categories were examined: a) ±3 days around the index prescription date, and b) 4 to 365 days before the index prescription date. To identify the diagnostic codes that mapped to each of the 13 antidepressant treatment indication categories, we used the list of ICD-9 codes from our previous study (148) (see Appendix A). ICD-10 codes recorded in hospital discharge abstracts from April 2006 onwards were back-translated to their ICD-9 equivalent using conversion tables.(134) We also created another 17 binary variables to capture the presence of diagnostic codes for conditions in the Charlson comorbidity index recorded over the past year since patients with depression often have more chronic morbidities than patients without depression.(151,152) Diagnostic codes for these conditions were identified using the ICD mappings by Quan *et al.*(127) 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 (see Appendix B, Table B-1).

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.(152) 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 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 – see Appendix B, Table B-2). 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 back-translated to their CCP equivalent using conversion tables from RAMQ. Binary variables were used to represent the presence of any 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 – see Appendix B, Table B-3).

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 (e.g. aripiprazole, quetiapine) or lithium, which are drugs commonly used to augment antidepressant therapy for depression.(153) Binary variables were used to represent the presence of previous prescriptions for any drug (molecule) that had been prescribed in the past year for at least 1% of patients with antidepressant prescriptions (a total of 99 drugs – see Appendix B, Table B-4).

7.5.6.3 Physician factors

Various factors may influence why physicians prescribe antidepressants including their previous medical training, clinical experience, workload, and receptiveness to information on best practices. To test these hypotheses, we used information from RAMQ to capture physicians' sex, place of medical training (Canada/US or other), level of clinical experience (expressed as the number of years in practice since medical graduation), and workload (expressed as the average number of patients seen per working day in the previous year). MOXXI physicians also completed the Evidence-Nonconformity-Practicality survey,(85) 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 to 30) indicate a stronger belief in scientific evidence versus clinical experience as the best source of clinical knowledge, higher nonconformity scores (possible range 6 to 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.

7.5.7 Statistical analysis

Figure 7-1 illustrates the steps of the study analysis. We used multivariable binomial logistic regression to model the probability that a given prescription was written for an indication besides 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 7-1). 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 included in the analysis and 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 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 7-1, Step 1). These blocks were used for 3-fold cross-validation to reduce the risk of overfitting the final model to the training set (154). 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 7-1, Step 2). We repeated this process three times, each time using a different validation set and then averaged the performance over the three validation sets. As the performance metric, we used the scaled Brier score (155,156), similar to the R^2 statistic in linear regression, which was calculated using the following formula:

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

where N represents the total number of antidepressant prescriptions, $\hat{Y}_i$ represents the predicted probability that prescription i was written for an indication besides 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 besides 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 besides 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 7-1, Step 2A). To this end, we relied on the flexible yet parsimonious first-degree fractional polynomials (FP1) (157). For each continuous variable *X*, we selected the best fitting FP1 function among eight candidate FP1 functions: $\beta_1 X^p$, where the powers *p* were represented by the set {-2, -1, -0.5, 0, 0.5, 1, 2, 3}, and X^0denoted log(*X*) (157). The best functional form for each *X* was assumed to be X^1 (i.e. a linear association with outcome) unless a model using one of the non-linear FP1 functions had a cross-validated scaled Brier score that was at least 0.0005 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 7-1, 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 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 minimal improvement in predictive performance). Finally, among the covariates added to the model from the 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 (i.e. without interactions) by at least 0.0005 (Figure 7-1, 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 7-1, Step 3). We then used the fitted coefficients of these models to predict the outcome for prescriptions in the test set and evaluated their performance using the methods and criteria described below (Figure 7-1, Step 4).

7.5.7.1 Measures of predictive performance

7.5.7.1.1 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.(156) 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 besides diagnostic codes for plausible antidepressant treatment indications in the prediction models .(158) 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 (i.e. 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.(158)

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 besides 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 besides depression in each stratum was calculated by summing the estimated probabilities across all prescriptions in the corresponding stratum.

7.5.7.1.2 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:

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

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 (i.e. 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. (155) 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 (103) 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 2.5^{th} and 97.5^{th} percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples of the test set.

7.5.7.2 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 7-1, 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 (103) to calculate the 95% CIs around the OR_{adj} estimates.

All statistical analyses were performed using SAS (SAS Institute) software, version 9.4.

7.6 Results

The study analysis included a total of 73 576 antidepressant prescriptions written by 141 physicians for 16 262 patients (Figure 7-1). 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 besides depression – 16 374 (22.3%) for anxiety/stress disorders, 7295 (9.9%) for sleeping disorders, 4137 (5.6%) for pain, and 4599 (6.3%) for miscellaneous indications. The remaining 41 171 (56.0%) prescriptions were for depression.

7.6.1 Prescription, patient, and physician characteristics

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.

Approximately two-thirds (67.4%) of patients were female. At the time of the antidepressant prescription, 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 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 (92.2%) physicians received their medical training in Canada or the US and most (77.3%) had been practicing for at least 15 years at the start of the study. Physicians saw a median of 7 (IQR 13-19) patients per working day. Physician scores on the Evidence-Nonconformity-Practicality survey (85) suggested that they favored scientific evidence over clinical experience as the best source of knowledge (median evidence score of 21, IQR 20-23), 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 13-17).

7.6.2 Best FP1 function for continuous variables

Among all candidate predictors considered, 13 were continuous variables. For seven of these covariates, there was a better non-linear 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 7-1).

7.6.3 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 7-2). 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 besides 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 7-2). 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 besides depression. Adding this variable resulted in a cross-validated scaled Brier score of 0.3193 – 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 (e.g. 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 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 7-2).

7.6.4 Performance of the final versus baseline models

7.6.4.1 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 7-3). 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 7-3) 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 besides depression (Table 7-4). 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 (0.2-0.6). 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 besides depression. Furthermore, the baseline model did not afford a clear identification of prescriptions with a high probability of indications besides 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 7-4).

7.6.4.2 Secondary outcome

For the multinomial logistic regression model that predicted a specific class of antidepressant treatment indications, the scaled Brier score across the five treatment indications classes was 0.320 (95% CI 0.249 to 0.385) (Table 7-5). 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 the miscellaneous indication category (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 0.029, 95% CI 0.004 to 0.043).

7.6.5 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 be prescribed for indications besides 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 7-6). 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 besides 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 besides 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 7-6). On the other hand, antidepressants that were prescribed on a 'take-as-needed' basis were notably more likely to be prescribed for indications besides 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 besides depression, with these associations being strongest when the diagnostic code was recorded around the index prescription date. On the contrary, patients were significantly less likely to be prescribed antidepressants for indications besides depression if they had a diagnostic code for depression, unspecified nonpsychotic mental disorder following organic brain damage (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 besides depression, but these 95% CIs did not exclude 1 (Table 7-6).

Figure 7-2 shows the adjusted ORs for the primary outcome that were estimated in the final model for the three continuous covariates that were expressed using non-linear 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 the values above a threshold.

7.7 Discussion

7.7.1 Main findings

In this study, we derived and validated a model that could accurately predict when antidepressants were prescribed for indications besides 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.

7.7.2 Comparison with previous studies

Few studies have attempted to predict antidepressant treatment indications or identify factors associated with them. Gardarsdottir *et al.* (59) 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.* (31) 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 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.* (18) identified factors associated with non-psychiatric antidepressant use in an adult Finnish population where patients' psychiatric history was assessed via self-report in a structured interview. The authors' definition of psychiatric use included both depression and anxiety disorders, but they similarly

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 association was statistically significant).

In all these studies, treatment indications were measured using suboptimal methods. Relying on diagnostic codes alone identifies the true indication for only a small proportion of antidepressant prescriptions (148) and using self-reported psychiatric history may lead to under-reporting of depression and other mental conditions.(18) 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.

7.7.3 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 (159) and chronic pain (160) 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.(161) 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.(162) 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 (163) 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 (e.g. pain, fibromyalgia, Crohn's disease) may require patients to undergo various diagnostic tests before reaching a proper diagnosis.(164,165) 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.

7.7.4 Implications of findings

Administrative data and electronic health records rarely contain information linking prescriptions directly to their corresponding diagnoses,(29,31) 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 besides depression is a positive finding for health services researchers. In the absence of documented treatment indications, our predictive model may allow researchers to still stratify antidepressant users by indication, thus representing a valuable tool for enabling more accurate database research on antidepressants. 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.

7.7.5 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 prevent 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 (18,31,59) 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.

Study limitations include the generalizability of our findings, as the MOXXI system is used by physicians in one Canadian province for patients who are generally older with more health complexities.(80) We did not consider drugs previously dispensed to patients because dispensing data was unavailable for nearly half (44%) of antidepressant prescriptions in this study where the patient had private drug insurance. Another study limitation is that treatment indications were modelled 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, which was similar to a UK study (16) that found only 5% of antidepressants users reported taking antidepressants for multiple indications. Furthermore, MOXXI physicians have good reason to enter all relevant indications into the MOXXI system because these conditions become part of the patient's problem list, which is then accessed by a drug knowledge databases to identify potential drug-disease problems.(36)

Finally, two study considerations deserve mention. First, because we used prescription data to derive our prediction model, its performance on prescription claims data may differ if the

characteristics of patients who fill their antidepressant prescriptions are distinct from the characteristics of 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.

7.7.6 Conclusions

In this study, we used administrative health data to derive a model that could accurately identify when antidepressants in primary care were prescribed for indications besides depression. In the absence of documented treatment indications, this model represents a valuable tool that could enable health services researchers to conduct more accurate database studies on antidepressant use by indication.

Table 7-1 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	
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}

Abbreviations: FP1 = first-degree fractional polynomial

[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 $p \leq 0$ 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 the Appendix.
[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*. (127)
[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 3 levels: 1) 24+ years, 2) 15 – 23 years, or 3) <15 years.
[n]Measured using physician scores on the Evidence-Nonconformity-Practicality survey,(85) 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.

Table 7-2. Derivation of the final prediction model

Order added	Variables included in the model	Scaled Brier score$_{CV}$[a]	Δ[b]
-	26 binary variables for the presence of diagnostic codes for 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

[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.

Table 7-3. Performance of the final and baseline models for predicting antidepressant prescriptions for indications besides depression

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

Abbreviations: IDI = integrated discrimination improvement

[a]Based on the coefficients for the final and baseline prediction models that were fit using the entire training set.

[b]Similar to the R^2 statistic for linear regression, where higher scores indicate better performance.

[c]Calculated as the absolute difference in the average estimated probability of indications besides depression among prescriptions for depression and not for depression.

[d]Quantifies the incremental value of adding other health-related information to the baseline model that used only information on diagnostic codes for plausible antidepressant treatment indications. The IDI is equal to the difference in discrimination slopes between the final and baseline models (158).

Table 7-4. Calibration of the final and baseline models for predicting antidepressant prescriptions for indications besides depression

Probability of treatment indication besides 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	5531	756	571.15	1.324 (0.911 to 1.701)	2398	424	385.67	1.099 (0.825 to 1.514)
>0.2 – 0.4	5646	1551	1703.99	0.910 (0.659 to 1.215)	3497	936	1070.39	0.874 (0.733 to 1.054)
>0.4 – 0.6	4427	2088	2171.00	0.962 (0.730 to 1.243)	11538	5331	5765.59	0.925 (0.736 to 1.158)
>0.6 – 0.8	2254	1521	1544.70	0.985 (0.833 to 1.152)	4049	2515	2624.67	0.958 (0.792 to 1.134)
>0.8 – 1.0	3699	3357	3417.19	0.982 (0.931 to 1.030)	75	67	66.11	1.013 (0.768 to 1.133)
Overall	21 557	9273	9408.03	0.986 (0.842 to 1.136)	21 557	9 273	9912.43	0.935 (0.773 to 1.125)

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

[a]The probability of the outcome was calculated for prescriptions in the test set based on the coefficients for the final and baseline models that were obtained using the entire training set.

[b]The expected number of antidepressant prescriptions with a treatment indication besides depression was calculated by summing the estimated probabilities across all prescriptions in the stratum.

Table 7-5. Overall and per-class performance of the final and baseline models for predicting antidepressant treatment indications expressed as a five-class outcome

	Performance in the test set	
	Scaled Brier score[a] (95% CI)	
Treatment indication class[b]	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 indications	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)

[a]Based on coefficients for the final and baseline models that were fit using the entire training set.
[b]The per-class estimate for each treatment indication category was calculated using a 'one-versus-rest' approach.

Table 7-6. Independent association between variables in the final prediction model and antidepressant prescriptions for treatment indications besides depression

	N[a] (%) or Median (IQR)	Antidepressant prescriptions for indications besides depression	
		Adjusted OR[b]	95% CI[c]
Prescription-related factors			
Molecule name			
Venlafaxine[e]	15 398 (20.9)	1.00	[Reference]
Amitriptyline[f]	6 196 (8.4)	20.98	12.27 to 48.91
Trazodone	6 891 (9.4)	18.55	8.7 to 45.88
Nortriptyline[f]	434 (0.6)	16.32	5.43 to 190.13
Doxepin[f]	461 (0.6)	10.49	2.60 to 109.35
Imipramine[f]	200 (0.3)	8.84	1.39 to 301.56
Desipramine[f]	138 (0.2)	3.98	1.31 to 73.55
Duloxetine[e]	1 596 (2.2)	2.40	1.10 to 6.10
Paroxetine[d]	6 751 (9.2)	2.05	1.11 to 3.64
Clomipramine[f]	165 (0.2)	1.54	0.26 to 12.36
Citalopram[d]	13 623 (18.5)	1.07	0.67 to 1.69
Escitalopram[d]	4 470 (6.1)	0.82	0.53 to 1.51
Sertraline[d]	4 457 (6.1)	0.74	0.45 to 1.26
Fluvoxamine[d]	669 (0.9)	0.72	0.20 to 1.50
Trimipramine[f]	436 (0.6)	0.69	0.20 to 3.01
Fluoxetine[d]	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
Desvenlafaxine[e]	477 (0.7)	0.18	0.02 to 310670.11

Abbreviations: OR = odds ratio, IQR = interquartile range

[a]N total prescriptions = 73 576

[b]Adjusted ORs were obtained using coefficients from a multivariable logistic regression model that was fit using all prescriptions (i.e. training and test sets combined).

Table 7-6 (continued). Independent association between variables in the final model and antidepressant prescriptions for treatment indications besides depression

	N[a] (%) or Median (IQR)	Antidepressant prescriptions for indications besides depression	
		Adjusted OR[b]	95% CI[c]
Prescription-related factors (continued)			
Prescribed dose (mg/day), per 10 mg increase by molecule			
Mirtazapine	30 (15 – 30)	0.68	0.48 to 0.89
Nortriptyline[f]	25 (10 – 50)	0.68	0.30 to 0.92
Paroxetine[d]	20 (15 – 30)	0.78	0.66 to 0.91
Desvenlafaxine[e]	50 (50 – 100)	0.83	0.05 to 1.13
Doxepin[f]	40 (25 – 75)	0.85	0.62 to 1.07
Citalopram[d]	20 (20 – 30)	0.86	0.76 to 0.96
Imipramine[f]	50 (25 – 75)	0.86	0.62 to 1.11
Fluoxetine[d]	20 (20 – 40)	0.87	0.65 to 1.11
Desipramine[f]	50 (25 – 100)	0.90	0.52 to 1.07
Amitriptyline[f]	20 (10 – 30)	0.92	0.78 to 0.99
Escitalopram[d]	10 (10 – 20)	0.95	0.64 to 1.14
Duloxetine[e]	60 (30 – 60)	0.96	0.80 to 1.11
Venlafaxine[e]	75 (75 – 150)	0.96	0.94 to 0.98
Trimipramine[f]	50 (25 – 75)	0.96	0.64 to 1.10
Sertraline[d]	50 (50 – 100)	0.99	0.94 to 1.02
Trazodone	50 (50 – 100)	0.99	0.95 to 1.05
Fluvoxamine[d]	100 (50 – 143)	1.00	0.90 to 1.09
Clomipramine[f]	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

Abbreviations: OR = odds ratio, IQR = interquartile range

[a]N total prescriptions = 73 576

[b]Adjusted ORs were obtained using coefficients from a multivariable logistic regression model that was fit using all prescriptions (i.e. training and test sets combined).

Table 7-6 (continued). Independent association between variables in the final model and antidepressant prescriptions for treatment indications besides depression

	N[a] (%) or Median (IQR)		Antidepressant prescriptions for treatment indications besides depression			
			Adjusted OR[b]		95% CI[c]	
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 to 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	
Antidepressant treatment indications	*± 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	3881 (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 91807.54	1.41	0.49 to 3.75
Pre-menstrual dysphoric disorder	26 (0.0)	82 (0.1)	1.77	0.22 to 80903.71	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	

Abbreviations: OR = odds ratio, IQR = interquartile range

[a]N total prescriptions = 73 576

[b]Adjusted ORs were obtained using coefficients from a multivariable logistic regression model that was fit using all prescriptions (i.e. training and test sets combined).

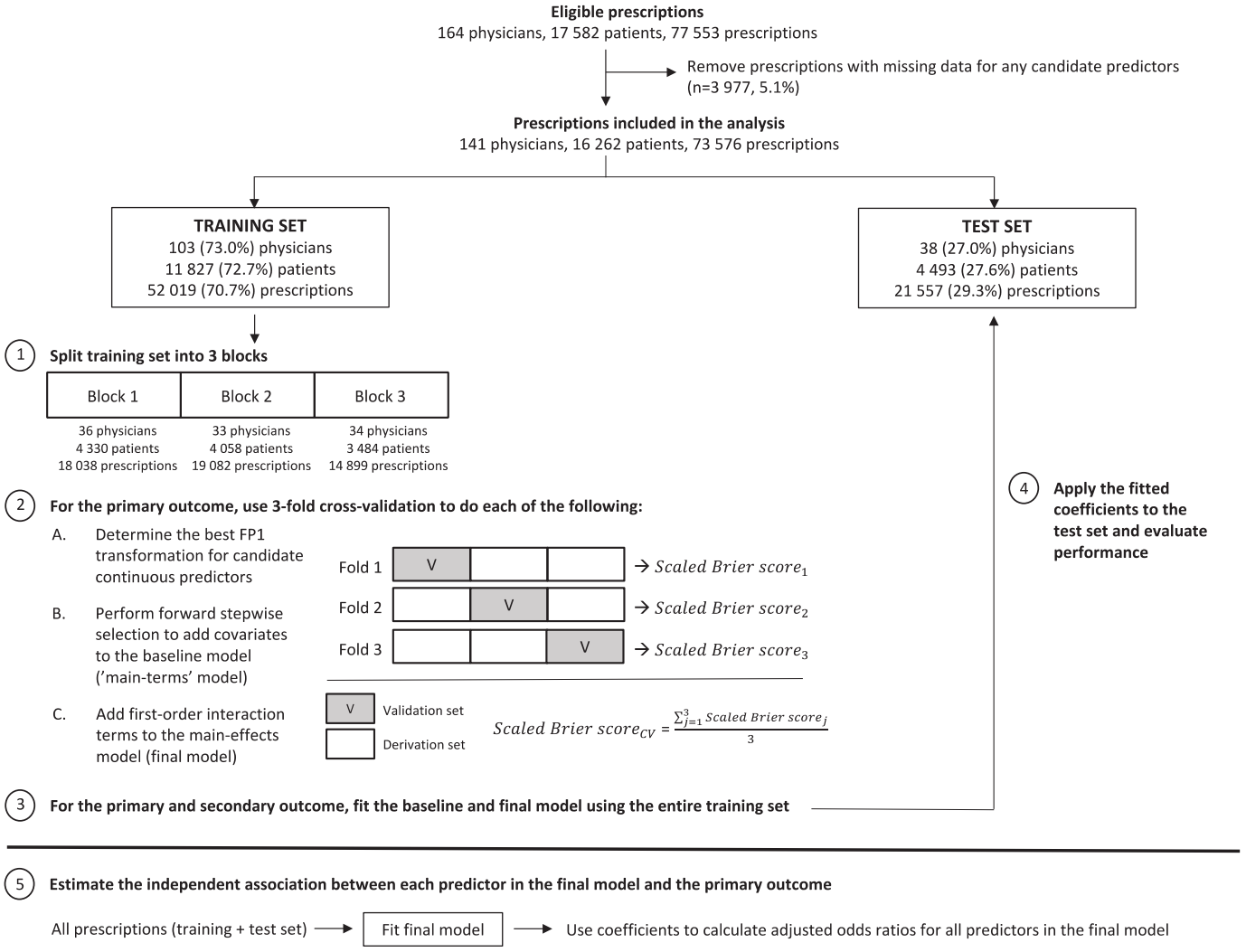


Figure 7-1. Flowchart of the study analysis

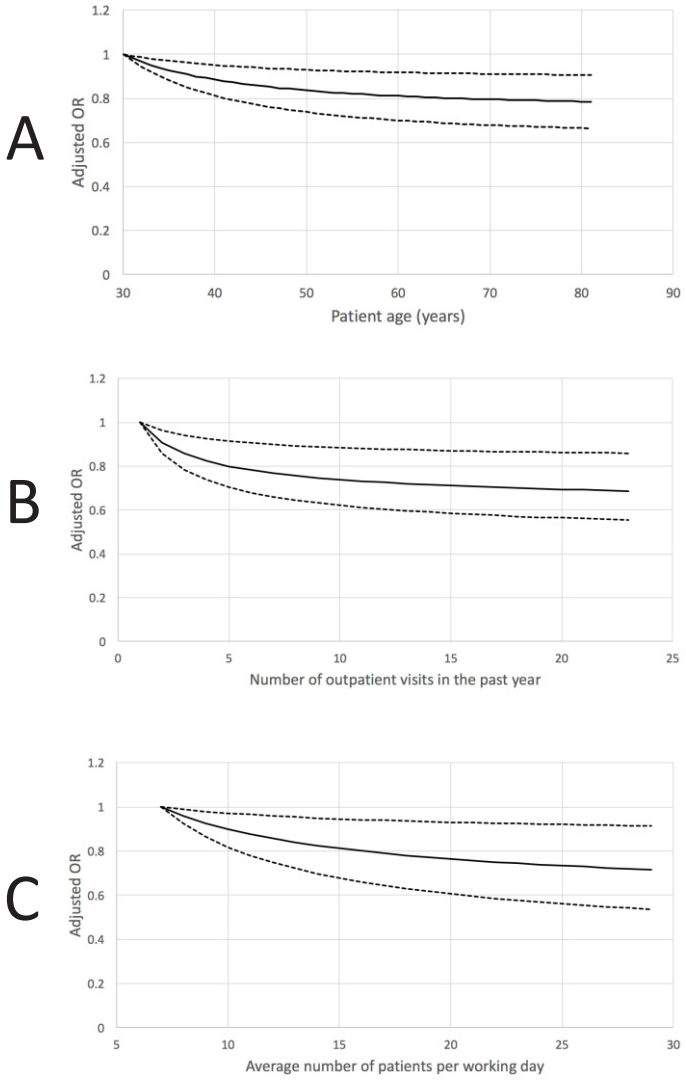


Figure 7-2. Independent association between antidepressant prescriptions for indications besides depression and the three continuous covariates in the final model that were expressed using non-linear FP1 functions

Abbreviations: OR = odds ratio. 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.

8 Optimizing the hyperparameter values of machine-learning algorithms improved the performance of super learning for predicting when antidepressants were prescribed for indications besides depression

8.1 Preamble

In manuscript 4, I used standard regression techniques to predict when antidepressants were prescribed for indications besides depression. Given the complexity of the dataset I used and the large number of covariates that I considered, the use of more contemporary methods for this prediction task may achieve even better performance than standard regression methods. Thus, in this manuscript, I explored the use of more flexible machine-learning algorithms and the ensemble learning approach called “super learning” to predict when antidepressants were prescribed for indications besides depression. I also determined whether optimizing the hyperparameter values of machine-learning algorithms improved the performance of super learning since many previous studies in the epidemiological literature do not appear to have optimized hyperparameter values when using super learning.

8.2 Title page and footnotes

Title: Optimizing the hyperparameter values of machine-learning algorithms improved the performance of super learning for predicting when antidepressants were prescribed for indications besides depression

Authors: Jenna Wong[1], MSc; Travis Manderson, MSc[2]; Michal Abrahamowicz[1], PhD; David Buckeridge[1], MD, PhD; and Robyn Tamblyn[1], PhD

Affiliations

[1]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

[2]School of Computer Science, McGill University, Montreal, Canada

Corresponding author:

Jenna Wong

1140 Pine Avenue West

Montreal, Quebec, Canada

H3A 1A3

Email: jenna.wong@mail.mcgill.ca

8.3 Abstract

Background: Super learning is an ensemble machine-learning method that is being increasingly used to predict outcomes and create propensity scores for the estimation of causal effects. However, the hyperparameters of machine-learning algorithms may not always be considered by investigators when using super learning. This study determined whether optimizing the hyperparameter values of machine-learning algorithms improved the performance of super learning.

Methods: We used super learning to predict when antidepressant prescriptions were written for indications besides depression. Our analytical dataset for this prediction task included 73 576 antidepressant prescriptions and over 370 variables. We considered five popular machine-learning algorithms in the super learner and used an iterative grid search procedure to identify the optimal value for their hyperparameters. We compared the performance of a super learner using the optimal hyperparameter values to the performance of a super learner using the default values in the algorithms' respective statistical packages.

Results: Overall, 44% of antidepressant prescriptions were written for indications besides depression. The super learner using the optimal hyperparameter values outperformed the super learner using the default values by 4% (95% CI 1%–8%). The discrimination of the super learner was also better when using the optimal hyperparameter values instead of the default values (*c* statistic of 0.822, 95% CI 0.795-0.847 compared to 0.817, 95% CI 0.791-0.846).

Conclusion: Super learning is a powerful method for combining the predictions from machine-learning algorithms, but to achieve the best performance, investigators should optimize the hyperparameter values of algorithms in the super learner.

Keywords (3-10): super learning, hyperparameter optimization, grid search, machine-learning, predictive modelling

8.4 Background and significance

Predictive modeling has many important applications in public health, clinical practice, and epidemiological research. Risk scoring systems and prediction algorithms can help policy makers target public health interventions to high-risk populations, enable physicians and patients to make more informed treatment decisions based on the probability of disease or prognostic outcomes, and help control for confounding in observational studies through the creation of propensity scores (166). The rise of Big Data in healthcare has the potential to greatly improve our ability to accurately predict outcomes (167), but it is no trivial task sorting through the masses of data to separate the signal from the noise.

Standard practices for deriving risk prediction models typically involve using parametric regression methods (e.g. linear or logistic regression) where the optimal set of covariates is identified by using a stepwise variable selection procedure (e.g. forward selection, backward elimination), testing for interactions, and trying different functional forms for continuous variables (e.g. polynomials, splines). Such practices are important because the probability estimates from regression models may be biased if the model is incorrectly specified (38). However, as the dimensionality (i.e. number of covariates) of the dataset grows, researchers may find that these standard procedures become cumbersome and difficult to implement properly.

Because of these challenges, there has been growing interest in the medical and epidemiological community to use more flexible prediction techniques from the machine-learning literature that can automatically learn associations from data (168–170). For example, neural nets and support vector machines (SVMs) are well-suited to handle high-dimensional data and can capture complex interactions and functions without requiring the investigator to explicitly add extra interaction or polynomial terms to the model (168). Algorithms like decision trees, random forests, and penalized regression automatically perform variable selection as part of the algorithm's learning (optimization) process (154).

Given these advantages, the potential for other machine-learning algorithms to outperform standard regression techniques while requiring less input from the investigator is an enticing reason to explore their use. Because there are many possible algorithms from which to choose, the use of "super learning" (39) has become a popular approach for predicting outcomes in epidemiology. Super learning is a machine-learning ensemble method that fits a series of user-specified algorithms and combines them using an optimal, weighted linear combination of their predictions (39).

Super learning can be implemented using packages like *SuperLearner* (40) in the R programming language. Because of the "black box" nature of this package (171), it does not appear that users often modify the default values of the hyperparameters for algorithms included in the super learner (41–45,171). These hyperparameters, however, affect the performance of machine-learning algorithms by controlling their complexity (46,154). To obtain the best performance of an algorithm for a specific task, its hyperparameters should be tuned to their optimal values. This tuning is often done through an iterative process called *grid search* whereby the algorithm is repeatedly assessed at different possible hyperparameter values according to a cross-validated performance metric (46,169). Thus, when machine-learning algorithms are run "as is" (i.e. using the default hyperparameter values in their respective statistical packages), the algorithms – and therefore the super learner – may not perform as well as if the hyperparameters had been properly tuned.

To test this hypothesis, we compared the performance of super learning when the hyperparameters of its algorithms were tuned using grid search versus set to the default values in their respective statistical packages. For this analysis, we employed the same dataset from our previous study (172) that used an extensive number of variables derived from health services data to predict when antidepressants were prescribed for indications besides depression. The ability to accurately identify patients who are not taking antidepressants for depression is important because treatment indications for medications are not routinely recorded, thus creating major barriers to evaluating antidepressant use for off-label indications.

In our previous analysis (172), we only applied standard regression techniques for prediction. In this study, we explore the use of machine-learning algorithms and super learning to predict antidepressant treatment indications.

8.5 Methods

8.5.1 Data source

The Medical Office of the XXIst Century (MOXXI) is an indication-based electronic prescribing and drug management system that is used by over 185 consenting primary care physicians at community-based clinics around two major urban centers in the Canadian province of Quebec (36). An important feature of the MOXXI system is that physicians are required to document at least one treatment indication for every prescription 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. Treatment indications in the MOXXI system were previously validated against a blinded, post-hoc physician-facilitated chart review where they had excellent sensitivity (98.5%) and high positive predictive value (97.0%) (37). Health services data on all patients in the MOXXI system are also available through the system's integration with the provincial health insurance agency (The Régie de l'assurance maladie du Québec [RAMQ]) and the provincial hospital discharge summary database (MED-ECHO). Together, these data sources provide information on patient demographics, diagnoses, hospitalizations, and medical services received.

This study included all drugs approved for depression that had at least 120 prescriptions written in the MOXXI system between 1 January 2003 and 31 December 2012. The unit of analysis was the antidepressant prescription. All patients gave informed consent to have their information used for research purposes. This study was approved by the McGill institutional review board.

8.5.2 Study variables

The outcome being predicted was a binary variable indicating whether an antidepressant prescription had been prescribed for an indication besides depression. The outcome was measured using the physician-documented treatment indications in the MOXXI system.

Table 8-1 lists all covariates that were included in the analysis. There were a total of 373 variables related to characteristics of the prescription, patient, or prescribing physician. Prescription-related variables (n=4) included the molecule name, the prescribed dose, whether the drug was supposed to be taken on an 'as-needed' basis, and the number of other drugs concurrently prescribed with the antidepressant. Patient-related variables (n=362) captured information on demographics, socio-economic status, diagnostic codes, health services use (e.g. previous hospitalizations, outpatient visits, ER visits, medical services received), and drugs prescribed in the past year. Finally, physician-related variables (n=7) included physician sex, place of medical training, level of clinical experience, size of patient workload, and scores from a survey (85) that measured physicians' attitudes towards new information about good clinical practice. Details about the creation of these variables were described in the earlier paper (172).

Of the 373 covariates, 13 were continuous variables, two were categorical variables, and the remaining were binary variables. Each categorical variable was expressed using a series of *n-1* binary variables, yielding a final covariate matrix with 391 columns. Because some machine-learning algorithms (e.g. neural nets, SVMs) require scaling of the inputs, each continuous variable was standardized by subtracting the mean of its values and dividing by twice the standard deviation (154).

8.5.3 Machine-learning algorithms considered and their hyperparameters

We considered five machine-learning algorithms in the super learner: 1) least absolute shrinkage and selection operator (LASSO) penalized regression, 2) recursive partitioning and regression trees (hereafter referred to as decision trees), 3) random forests, 4) neural networks, and 5) SVMs (Table 2). We chose these algorithms because they used a broad range of

approaches for solving prediction tasks and for their popularity in fields like genetics (173) and biomedicine (174). We limited the super learner to a collection of five algorithms because of the computational resources required to optimize each of their hyperparameters. Table 8-2 shows the R packages we used to implement each of the machine-learning algorithms and the hyperparameters that we optimized.

LASSO penalized regression simultaneously (i) shrinks the coefficients of a conventional regression model towards zero and (ii) performs variable selection by shrinking some of the coefficients right to zero (154). Coefficients with higher variance are shrunk more. The amount of shrinkage increases as the value of regularization parameter lambda increase, and a lambda of zero yields the conventional unpenalized logistic regression model.

Decision trees are non-parametric learning algorithms that apply a set of rules to partition the multidimensional space of covariates into hypercubes within which the outcome is fairly homogeneous (173). Although decision trees are powerful and well-suited to handle high-dimensional data, they often produce results that are numerically unstable (i.e. slight changes in the data can produce notably different trees) and are prone to overfitting as the depth (or complexity) of the tree increases (173). To reduce overfitting, a stopping rule is often applied (154). In the R package *rpart* (175), this stopping rule is controlled by the hyperparameter cp (which stands for "complexity parameter) that retains only those splits that improve the overall performance of the tree by a factor of cp. Thus, larger values of the hyperparameter cp imply smaller, simpler trees with fewer nodes.

Random forests is an ensemble learning method that extends the decision tree framework in an attempt to address the issues of overfitting and high variability (176). Rather than a single tree, random forests contain many trees (typically hundreds) – each grown on a separate bootstrap re-sample of the training data where a random subset of predictors is chosen as the candidates at each split (154). The key hyperparameters for random forests are the number of trees to

grow (ntree) and the number of predictors to randomly select for consideration at each node (mtry) (173).

Artificial neural networks, or simply neural nets, are non-linear statistical models that attempt to emulate the complex structure of the human brain (177). Neural nets consist of an input layer (i.e. the variables offered to the neural net), one or more "hidden" layers, and an output layer that yields the predicted probabilities from the network. Each layer in the network contains a certain number of units or nodes that are connected to nodes in the subsequent layer by "connection weights" that act much like the beta coefficients in a regression model (178). Each node takes a weighted linear combination of its inputs (i.e. the sum of its inputs multiplied by their connection weights) and passes this result through an "activation function" (usually the logistic or sigmoid function), which then becomes input to the node(s) in the next layer to which it is connected via another connection weight. These hidden nodes and their connection weights are what allow neural nets to model complex non-linear relationships (178). Using neural nets with two or more hidden layers is often referred to as "deep learning", but one hidden layer is often sufficient for most applications in epidemiology (168,178). In this study, we considered neural nets with one hidden layer and optimized the number of hidden nodes (size) in this layer.

Finally, SVMs are algorithms that classify observations by finding the optimal hyperplane, in the multidimensional space defined by the potential predictors, that separates observations of different outcome classes with the maximum 'margin' (the distance between the hyperplane and the nearest data points of different outcomes classes on either side of it, called the "support vectors") (154). In practice, it may be challenging to find a hyperplane that perfectly separates two classes of the outcome. Thus, SVMs allow for a "soft margin" whereby a fraction of the datapoints can be on the wrong side of the hyperplane. The regularization parameter C controls the trade-off between minimizing the number of misclassified examples on the wrong side of the hyperplane versus maximizing the margin by determining the strength of the penalty for misclassification (179). A smaller C generally favors a larger margin (i.e. smoother decision

surface) because the penalty for misclassifying data points is small, whereas a larger C implies a higher penalty and thus favors a smaller margin (i.e. more complex decision surface) with a lower percentage of misclassified data points. SVMs also use kernel functions to increase the dimensionality of the input space, which often allows a hyperplane to better separate the data points of different classes and achieve more complex, nonlinear decision boundaries in the space of the original, untransformed values of the covariates (154). In this study, we used a radial basis function (RBF) kernel – one of the most commonly used kernels for SVMs (180,181), and optimized the gamma parameter of the RBF kernel, which controls the spread of the decision boundary, with a higher gamma resulting in a more complex decision boundary.

8.5.4 Statistical analysis

Figure 8-1 describes the steps of the study analysis. As in the previous study (172), only antidepressant prescriptions with complete data for all covariates were used in the main analysis (~95% of all eligible prescriptions). All prescriptions with complete data were randomly divided into a 'training set' versus 'test set' using a 3:1 split. The training set was used to tune the hyperparameters of the machine-learning algorithms and estimate the coefficients in the super learner prediction function. The test set was used only to evaluate the performance of the final algorithms – it was not used in any part of the training process so that the final algorithms would be tested on unseen data. Because prescriptions were clustered within patients, who in turn were nested within physicians, we assigned a random sample of 75% of study physicians (rather than prescriptions) to the training set and the remaining 25% of physicians to the test set. Thus, all prescriptions from the same physician and patient were limited to the 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 physicians separately within each stratum.

We used the same randomization procedure to divide physicians in the training set into three mutually exclusive blocks of approximately equal size (Figure 8-1, Step 1). These blocks were

used to perform 3-fold cross-validation during the training of all algorithms to reduce the likelihood of overfitting the algorithms to the training set (154). We included all covariates as predictors in the algorithms. To tune the algorithms' hyperparameters, we used a grid search procedure that assessed the cross-validated performance of the algorithms iteratively over a wide range of plausible values (Table 8-2). For the LASSO penalized regression, rather than define our own subset of possible lambda values for the grid search, we used the sequence of values automatically chosen by the *glmnet* package. For algorithms with multiple hyperparameters, we assessed all possible combinations of their hyperparameter values. For example, for random forests, we assessed a total of 5 x 6 = 30 combinations. As the performance metric, we used the scaled Brier score (155,156), similar to the R^2 statistic in linear regression (176), which was calculated using the following formula:

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

where N represents the total number of antidepressant prescriptions, $\hat{Y}_i$ represents the predicted probability that prescription i was written for an indication besides 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 besides depression in the study sample. As the formula shows, the scaled Brier score can be interpreted as the proportion by which a given algorithm reduces the mean squared error of a non-informative algorithm where all prescriptions are assigned the overall marginal probability of having an indication besides depression. To compute the cross-validated scaled Brier score for a given algorithm, we fit the algorithm on two of the three training blocks (the "derivation set") and calculated the scaled Brier score in the remaining block (the "validation set"). We repeated this process three times using a different block as the validation set each time, and then averaged the scaled Brier score over the three validation sets (Figure 8-1, Step 2). The optimal value for each hyperparameter (or combination of hyperparameters) was defined as value that produced the best cross-validated Brier score in the training set.

Next, we used the *SuperLearner* package in R to fit two super learner prediction functions (Figure 8-1, Steps 3 to 5). For the first super learner, the five algorithms were fit using their optimal hyperparameter values that were identified from the grid search. To set the algorithm hyperparameters to their optimal values, we customized their learner functions by using the *create.Learner* function in the *SuperLearner* package or by creating our own wrapper function for the algorithm. For the second super learner, the five algorithms were fit using the default hyperparameters values in their corresponding wrapper functions in the *SuperLearner* package (for example, *SL.randomForest* for random forests). The LASSO penalized regression was the only algorithm where the optimal and default value of the hyperparameter lambda was guaranteed to be the same. The reason being that the *SL.glmnet* function in the *SuperLearner* package automatically used the value of *lambda* with the lowest cross-validated error. The weighted combination of the five algorithms in each super learner function was determined as follows. First, we used the same 3-fold cross-validation procedure to obtain predicted values for each algorithm in the validation set of each fold (Figure 8-1, Step 3). We then calculated the optimal convex combination of weights (i.e. a vector of non-negative weights that summed to 1) by regressing the observed outcome on a matrix with five columns containing the cross-validated predictions from each algorithm (Figure 8-1, Step 4). The vector of coefficients from this regression model corresponded to the optimal weights for combining the algorithms in the super learner prediction function. Finally, we refit the individual algorithms using the entire training set (Figure 8-1, Step 5). The predictions from these fitted algorithms, combined with their corresponding weights, comprised the super learner prediction function.

We assessed the performance of the two super learner functions by applying their fitted algorithms and corresponding weights to the prescriptions in the test set (Figure 8-1, Step 6). To compare their performance, we calculated the relative efficiency (RE) where RE = Brier $\text{score}_{\text{scaled}}(\text{SuperLearner}_{\text{optimal}})$/Brier $\text{score}_{\text{scaled}}(\text{SuperLearner}_{\text{default}})$. Thus, an RE > 1 indicated the proportional efficiency gain of the super learner when using the optimal hyperparameter values instead of the default values. We also calculated the RE separately for each algorithm in the super learner, where RE = Brier $\text{score}_{\text{scaled}}(\text{algorithm}_{\text{optimal}})$/Brier $\text{score}_{\text{scaled}}(\text{algorithm}_{\text{default}})$.

Finally, we assessed the discrimination of the two super learner models by constructing ROC curves and calculating the *c* statistic (156). We calculated 95% confidence intervals (CIs) around all performance measures using a two-stage cluster bootstrap (103) to account for multi-level clustering of prescriptions. The reported 95% CIs correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples of the test set. All analyses were performed in the R environment for statistical computing, version 3.4.1 (182). The following R packages were used: *glmnet* (183)*, rpart* (175)*, randomForest* (184)*, nnet* (185)*, e1071* (186)*, SuperLearner* (40)*,* and *AUC* (187).

8.6 Results

The analysis included a total of 73 576 antidepressant prescriptions that were written by 141 physicians for 16 262 patients (Figure 8-1). Of these, 52 019 (70.7%) antidepressant prescriptions were randomized to the training set, while the remaining prescriptions were assigned to the test set. Overall, 32 405 (44.0%) antidepressant prescriptions were written for indications besides depression, with this prevalence being similar between the training set (43.0%) and test set (44.5%).

Based on findings from the grid search procedure with cross-validation, the optimal hyperparameter value (or combination of hyperparameter values) was different from the default value for the random forests, neural net, and SVM (Table 8-3). For the random forests, although the optimal and default number of trees in the forest (nTree) was the same, the optimal number of variables randomly sampled as candidates at each split (mTry) was 50 compared to the default number of 19. For the neural net, the optimal number of hidden nodes (size) in the single hidden layer was one compared to the default of two hidden nodes. For the SVM, the optimal and default value of the regularization parameter C was the same, but the optimal value of the gamma parameter in the RBF kernel was 0.01 compared to the default value of 0.00256.

Table 8-4 shows the weights or coefficients for each algorithm in the super learner prediction function when the algorithms were fit using their optimal hyperparameter values and their default values. When the optimal hyperparameter values were used, the decision tree did not contribute at all towards the super learner prediction function (weight of 0) while the random forests contributed the most at 0.526. Among the remaining algorithms, the neural net and LASSO each contributed equally at 0.186 and 0.173, respectively, while the SVM had the lowest non-zero weight at 0.114. When the default hyperparameter values were used, this time the SVM did not contribute at all to the super learner prediction function. The LASSO and random forests contributed the most with a weight of 0.424 each, while the neural net had a weight of 0.106 and the decision tree had the lowest non-zero weight at 0.045.

When the super learner prediction functions were applied to the test set (i.e. the predictions from the individual algorithms were combined using their super learner weights), the overall performance of the super learner using the optimal hyperparameter values was better than the super learner using the default hyperparameter values, with a relative efficiency gain of 4% (95% CI 1% – 8%) in the scaled Brier score (Table 8-5). The ROC curves for the two super learner functions revealed that the discrimination of the super learner was slightly better when it used the optimal hyperparameter values rather than the default values (Figure 8-2). The area under the ROC curve, or *c* statistic, was 0.822 (95% CI 0.795 – 0.847) for the super learner using the optimal hyperparameter values compared to 0.817 (0.95% CI 0.791 to 0.846) for the super learner using the default values.

Among the individual machine-learning algorithms, the SVM, followed by the random forests had the best performance (highest scaled Brier score) when either the optimal or default hyperparameter values were used (Table 8-5). The decision tree, whose optimal and default values were the same, had the worst performance of all the algorithms. The relative efficiency gain for using the optimal hyperparameter values compared to the default values was highest for the neural net, which had a scaled Brier score of 0.299 (95% CI 0.239 – 0.345) using the optimal number of hidden nodes compared to 0.239 (95% CI 0.177 – 0.289) using the default

number. Thus, the relative efficiency gain was 25% (95% CI 15% – 41%). The overall performance of the random forests and SVM was also better when using the optimal hyperparameter values rather than the default values (Table 8-5). Finally, for the LASSO and decision tree, these algorithms performed the same in both super learners (RE of 1.00, 95% CI 1.00 – 1.00) because their optimal and default hyperparameter values were the same. Compared to the individual performance of each machine-learning algorithm, the use of super learning produced a super learner function that outperformed the individual algorithms (Table 8-5).

8.7 Discussion

In this study, we set out to determine whether the performance of the ensemble machine learning approach called "super learning" improved when the hyperparameters of its individual algorithms were tuned to their optimal values rather than set to the default values in their respective packages. We tested this hypothesis in a large dataset containing over 370 variables where the task was to predict when antidepressant prescriptions were written for indications besides depression. Both super learners performed well, but the super learner that used the optimal hyperparameter values performed best.

A growing number of researchers are using super learning to predict outcomes (42–44) or construct propensity scores to estimate casual effects in epidemiology (188–190). And understandably so – the advantages of super learning are that the user can select an array of sophisticated machine-learning algorithms and optimally combine them into an even more powerful prediction function. Although the performance of machine-learning algorithms is highly dependent on the value of their hyperparameters (46), many studies using super learning do not mention anything about hyperparameters in the text (41–43,45). When hyperparameter values have been mentioned, they are usually the default values in the algorithm's corresponding statistical package (176,188). The findings from this study suggest that if investigators optimize the hyperparameter values of their algorithms first, they may be able to achieve even better predictive performance with super learning.

There are several reasons why users may not have optimized the hyperparameter values in their super learner functions. First, the *SuperLearner* package is a "black box" that allows users to easily run complex machine-learning algorithms without having to know much about the algorithms themselves. To optimize algorithm hyperparameters, however, users must understand the algorithm's architecture, know the main hyperparameters that influence its performance, and identify a plausible range of hyperparameter values to test. Second, users may find it challenging to modify the "black box" to change the value of the algorithm hyperparameters. In this study, we found that the *create.Learner* function in the *SuperLearner* package was very helpful for creating custom learners – as long as the hyperparameter of interest was included as a modifiable parameter in the algorithm's wrapper function in *SuperLearner*. In cases where this requirement was not met (e.g. the gamma parameter for the svm), we created our own custom wrapper by copying the code of the original wrapper function (e.g. *SL.svm*), modifying the value of the hyperparameter directly in the code, and then creating a new name for the wrapper function (e.g. *SL.mySVM*). Third, users may not optimize hyperparameter values because the process of manually searching over a grid of possible values to identify the best one is labour-intensive and computationally expensive, especially for algorithms with long training times (e.g. neural nets and SVMs). Because of this barrier, we suggest that users pick a smaller number of diverse algorithms for the super learner and ensure that each algorithm is tuned to its best performance before running the super learner. A super learner with fewer algorithms that are optimally tuned may achieve better performance than a super learner with many untuned algorithms.

In a previous analysis (172), we attempted the same prediction task using standard regression techniques (e.g. forward stepwise selection, polynomial terms for continuous variables, and interaction terms between covariates). The final logistic regression model had a scaled Brier score of 0.307 (95% CI 0.245 – 0.360) and a *c* statistic of 0.815 (95% CI 0.787 to 0.847). Thus, this logistic regression model outperformed each of the machine-learning algorithms when its hyperparameters were not tuned. However, when these algorithms used their optimal

hyperparameter values, most of them performed similarly, if not slightly better, than the standard logistic regression model. These findings provide empirical evidence to suggest that a well-specified logistic regression model may achieve performance that is comparable to other more sophisticated machine-learning algorithms and may even outperform these algorithms if their hyperparameters have not been properly tuned. Both super learner functions, however, still outperformed the standard logistic regression model, demonstrating the power of this ensemble learning approach.

Finally, this study has several considerations. First, unlike other studies (42,43,176) that have used their entire dataset to perform 10-fold cross-validation of the super learner algorithm itself, we only assessed the performance of the two super learners in the test set. Thus, we only performed a single cross-fold validation of the two super learner algorithms. To use any of the prescriptions in the training set to validate the performance of the super learner algorithms would have potentially overestimated the performance of the super learner with the 'optimal' hyperparameter values (since these hyperparameter values had been selected using the prescriptions in the training set). Because we only used a single cross-fold validation, we calculated the variance around the performance estimates by applying the two super learners to 1000 bootstrap resamples of the test set. Second, although the grid search procedure that we used in this study is one of the most common methods for selecting hyperparameter values, the manual and iterative nature of this process makes it quite labour-intensive and requires a certain level of expertise in computing and machine-learning (46). Users may therefore want to explore newer methods that are being developed to automatically and more efficiently select hyperparameter values for a given machine-learning task (46). Lastly, when interpreting our findings, readers should keep in mind the properties of our analytical dataset such as the sample size, number of variables, and types of variables (i.e. mostly binary), as these factors may differentially affect the performance of various machine-learning algorithms (191).

In conclusion, the findings from this study provide empirical evidence to suggest that investigators may be able to achieve better performance with super learning if the

hyperparameters of the algorithms are tuned to their optimal values rather than set to their default values. Because it can be labour-intensive and computationally expensive to search for the best hyperparameter values, we suggest that users identify a smaller number of algorithms for super learning and focus on carefully optimizing the value of their hyperparameters.

Table 8-1. Covariates included in the analysis as predictors of antidepressant treatment indications (n=373)

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

[a]19 levels: venlafaxine, duloxetine, desvenlafaxine, citalopram, paroxetine, escitalopram, sertraline, fluoxetine, fluvoxamine, amitriptyline, doxepin, nortriptyline, trimipramine, imipramine, desipramine, clomipramine, trazodone, bupropion, or mirtazapine. The 19 levels were expressed using n-1 = 18 binary variables.
[b]Area-level measure representing the median household income ($CAD) in the patient's census tract area.
[c]Area-level measure representing the percentage of adults in the patient's census tract area with less than university education.
[d]Area-level measure representing the percentage of unemployed adults in the patient's census tract area.
[e]For each of the two observation windows, binary variables were used to indicate whether diagnostic codes in physician billings data or hospital discharge abstracts were recorded for each of the following 13 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.

[f]Binary variables were used to indicate whether diagnostic codes for any of the following 17 conditions in the Charlson comorbidity index were recorded over the past year in physician billings data or hospital discharge abstracts: 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.

[g]86 binary variables were used to represent each four-digit ICD-9 code that was recorded in physician billings data or hospital discharge abstracts for at least 1% of all antidepressant prescriptions in the past year (after excluding diagnostic codes for antidepressant treatment indications and Charlson conditions).

[h]Expressed as the percentage of all outpatient visits in the past year that were made to the prescribing physician.

[i]Based on billing codes recorded for the patient in physician billings 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).

[j]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).

[k]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).

[l]3 levels: 1) 24+ years, 2) 15 – 23 years, or 3) <15 years. The three levels were expressed using *n*-1 = 2 binary variables.

[m]Measured using physician scores on the Evidence-Nonconformity-Practicality survey (85), which is a psychometric instrument for determining how physicians would likely respond to new information about evidence-based clinical practice.

Table 8-2. Machine-learning algorithms and their hyperparameters

Algorithm	R package	Hyperparameter	Description of hyperparameter	Subset of values assessed in the grid search
LASSO penalized regression	*glmnet*	lambda	Regularization parameter	Sequence values automatically selected by *glmnet*
Decision tree	*rpart*	cp	Complexity parameter where splits that decrease the overall lack of fit by at least a factor of cp are retained	{0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, 0.00005, 0.00001, 0.000005, 0.000001}
Random forests	*randomForest*	nTree mTry	Number of trees to grow Number of variables randomly sampled as candidates at each split	{10, 100, 1000, 1500, 2000} {10, 25, 50, 100, 150, 200}
Neural net	*nnet*	size	Number of nodes in the hidden layer	{1, 2, 3, 4, 5}
SVM	*svm*	C gamma	Regularization term Parameter in the radial basis function kernel	{0.001, 0.01, 0.1, 1, 10} {0.001, 0.01, 0.1, 1, 10}

Abbreviations: LASSO = least absolute shrinkage and selection operator, SVM = support vector machine

Table 8-3. Optimal and default hyperparameter values for each algorithm

Algorithm	Hyperparameter	Optimal value[a]	Default value[b]
LASSO penalized regression	lambda[c]	0.001	0.001
Decision tree	cp	0.01	0.01
Random forests	nTree	1000	1000
	mTry	50	19[d]
Neural net	size	1	2
SVM	C	1	1
	gamma	0.01	0.00256[e]

Abbreviations: LASSO = least absolute shrinkage and selection operator, SVM = support vector machine
[a]The value that produced the best cross-validated scaled Brier score during the grid search procedure
[b]Default value in the algorithm's corresponding wrapper function in the *SuperLearner* package
[c]The optimal and default values of lambda were guaranteed to be the same because the *SL.glmnet* function in the *SuperLearner* package automatically used the value of lambda with the lowest cross-validated error
[d]Calculated by the function using the formula, floor(sqrt(ncol(x))), where ncol(x) = 391
[e]Calculated by the function using the formula, 1/ncol(x), where ncol(x) = 391

Table 8-4. Weights for the individual algorithms in the super learning functions

	Weight in the super learner function	
Algorithm	Optimal values for hyperparameters	Default values for hyperparameters
LASSO penalized regression	0.173	0.424
Decision tree	0.000	0.045
Random forests	0.526	0.424
Neural net	0.186	0.106
SVM	0.114	0.000

Table 8-5. Performance of the super learner functions and the individual algorithms when using the optimal and default hyperparameter values

	Scaled Brier score (95% CI)		
Algorithm	Optimal values for hyperparameters	Default values for hyperparameters	RE[a] (95% CI)
SuperLearner	0.322 (0.267 – 0.362)	0.309 (0.256 – 0.353)	1.04 (1.01 – 1.08)
LASSO penalized regression	0.287 (0.225 – 0.339)	0.287 (0.225 – 0.339)	1.00 (1.00 – 1.00)
Decision tree	0.226 (0.168 – 0.276)	0.226 (0.168 – 0.276)	1.00 (1.00 – 1.00)
Random forests	0.301 (0.251 – 0.341)	0.294 (0.284 – 0.329)	1.02 (1.00 – 1.05)
Neural net	0.299 (0.239 – 0.345)	0.239 (0.177 – 0.289)	1.25 (1.15 – 1.41)
SVM	0.310 (0.251 – 0.356)	0.300 (0.246 – 0.345)	1.03 (0.98 – 1.08)

Abbreviations: CI = confidence interval, RE = relative efficiency, LASSO = least absolute shrinkage and selection operator, SVM = support vector machine

[a]The relative efficiency of using the optimal hyperparameter value(s) for the algorithm versus the default value(s) from its wrapper function in the *SuperLearner* package. RE = Brier $score_{scaled}(algorithm_{optimal})$/Brier $score_{scaled}(algorithm_{default})$.

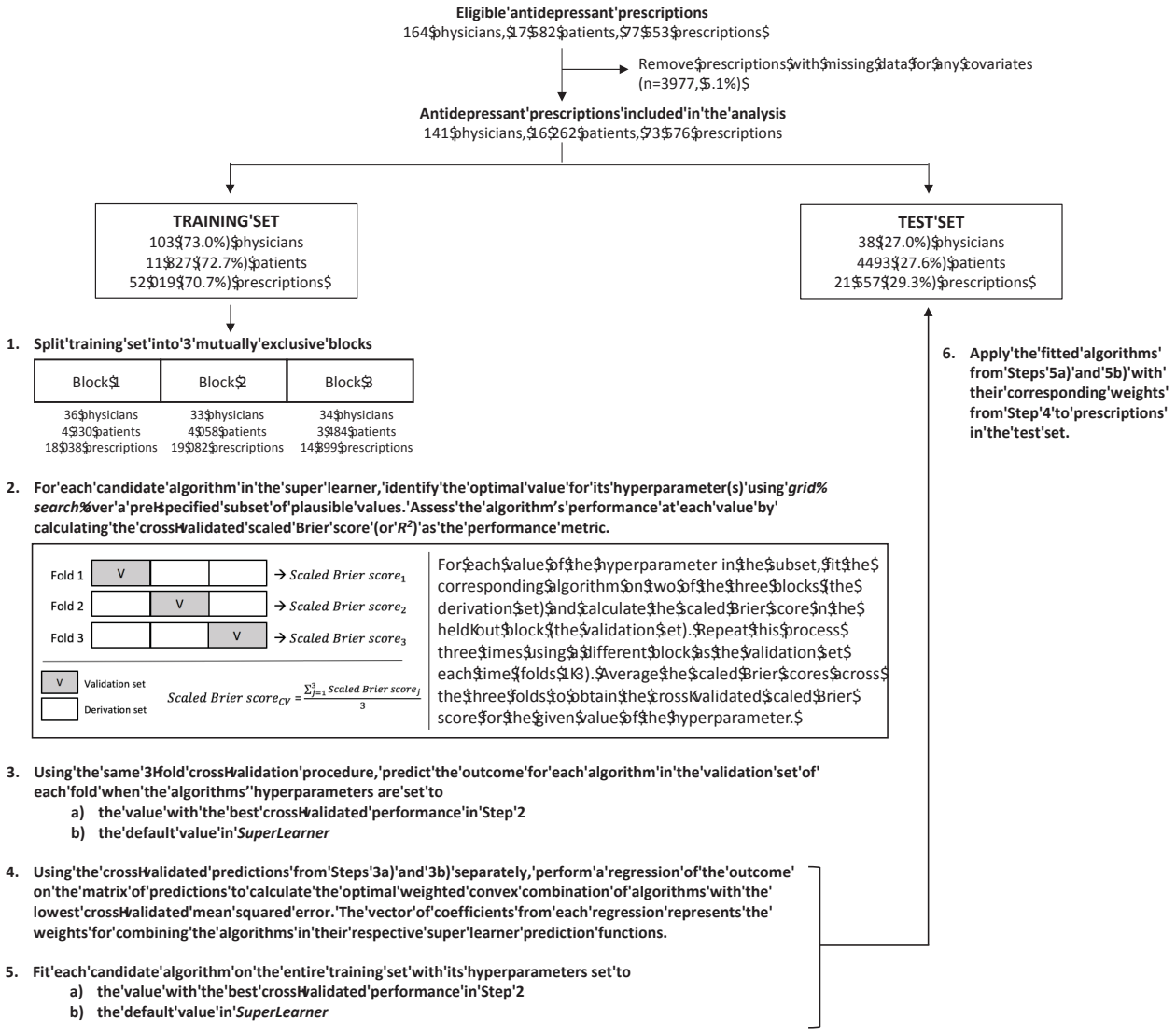


Figure 8-1. Flowchart of the study analysis

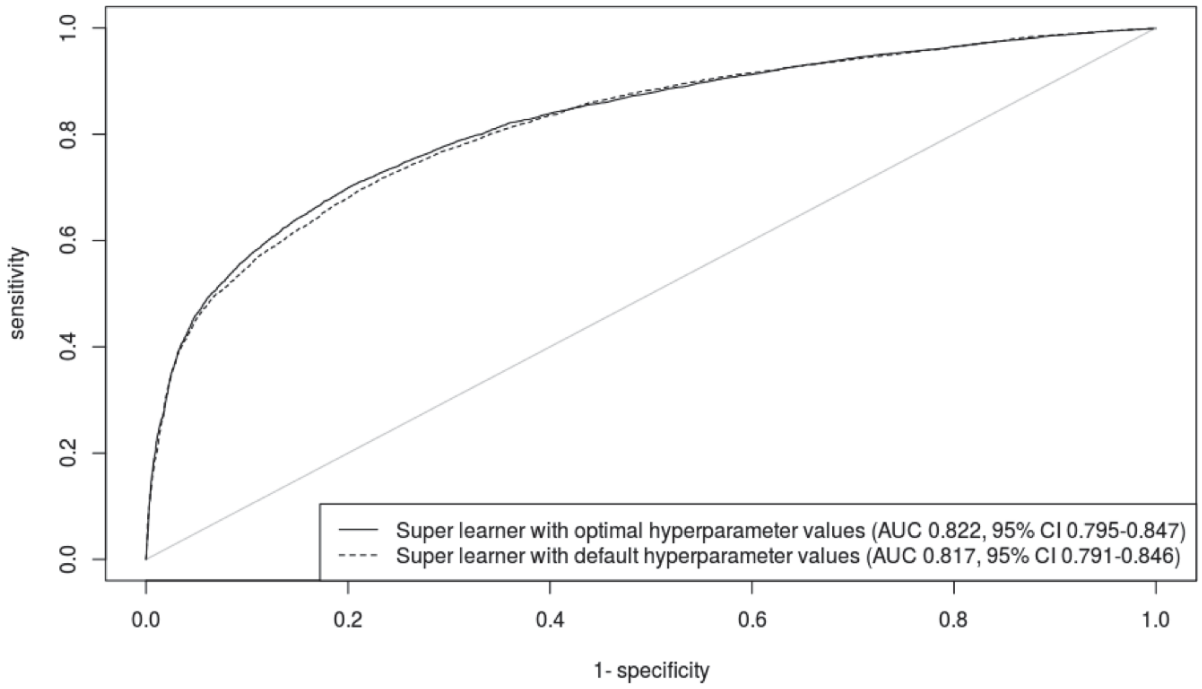


Figure 8-2. Receiver operating characteristic (ROC) curve for the two super learner functions

The solid black line shows the ROC curve for the super learner function when its algorithms were fit using the optimal hyperparameter values identified from the grid search. The dotted black line shows the ROC curve for the super learner function when its algorithms were fit using the default hyperparameter values in their respective functions within the *SuperLearner* package. Abbreviations: AUC = area under the receiver operating characteristic curve, CI = confidence interval.

9 Discussion

Antidepressants are among the most commonly used prescription drugs in North America (1,86). One of the factors driving their widespread use is a broadening of indications for these drugs, many of which are not licensed and have not been sufficiently evaluated. Given that antidepressants can cause various adverse side effects (24,25,192), the possibility that these drugs are being overprescribed for medical conditions that are unresponsive to treatment is a pharmacovigilance concern. Thus, it is important to monitor the use of antidepressants for non-evidence-based indications and to evaluate the outcomes associated with these uses. However, the ability to conduct these pharmacovigilance activities is severely hampered by the fact that treatment indications for medications are not routinely documented. Thus, the main goal of this thesis was to address these challenges and increase the capacity for comprehensive pharmacovigilance around antidepressants.

9.1 Summary of main findings

The four research objectives in this thesis were met through a series of five manuscripts, all of which used data from a unique electronic prescribing system that contained physician-documented treatments for all antidepressant prescriptions.

Objective 1 aimed to determine the prevalence of different treatment indications for antidepressant prescriptions, including off-label indications, and to assess the level of scientific support for these off-label indications. Manuscript 1 (Chapter 4) found that nearly half of antidepressant prescriptions in primary care were being written for indications besides depression, such as anxiety disorders, sleeping disorders, pain, and panic disorders, and that depression accounted for a decreasing proportion of antidepressant prescriptions over the past decade. I also found that one of every three antidepressant prescriptions was for an indication was that was not licensed for the drug. In Manuscript 2 (Chapter 5), the analysis showed that when antidepressants were prescribed for off-label indications, these off-label uses were usually not backed by sufficient scientific evidence. However, there was often another antidepressant in the same class as the prescribed drug that had strong evidence for the

respective indication. Together, the findings from these two manuscripts provided compelling evidence to indicate that more careful investigation is needed into the risks and benefits of off-label antidepressant use. Thus, the remaining objectives in this thesis addressed a major barrier to performing comprehensive pharmacovigilance for antidepressants, namely the ability to measure treatment indications for antidepressants.

Objective 2 aimed to assess the accuracy of using diagnostic codes from administrative health data to determine treatment indications for antidepressant prescriptions. This objective was met through manuscript 3 (Chapter 6), which considered 13 plausible indications for antidepressants and compared the presence of administrative diagnostic codes for these indications to the physician-documented treatment indications in the MOXXI system, which were considered the gold standard. Diagnostic codes for all 13 conditions were found to have poor sensitivity for identifying antidepressant prescriptions for the corresponding indication, especially among older patients and patients with more chronic comorbidities. These findings showed that the use of diagnostic codes alone is not sufficient to predict antidepressant treatment indications and could introduce substantial misclassification bias in analyses where this approach is used. Better methods for predicting treatment indications for antidepressants are therefore needed, which was the focus of the last two objectives in this thesis.

Objective 3 aimed to use standard regression techniques to predict antidepressant treatment indications based on a wealth of health services data beyond diagnostic codes. This objective was met through manuscript 4, which considered over 370 variables from health services data to derive and validate a logistic regression model that could accurately predict when antidepressant prescriptions were written for indications besides depression. This logistic regression model also identified important predictors of antidepressant treatment indications beyond diagnostic codes, including the molecule name, the patient's level of education, the physician's workload, and the dose of the prescribed drug.

Finally, objective 4 aimed to use more flexible machine-learning algorithms and "super learning" to predict antidepressant treatment indications. This objective was met through manuscript 5, which combined five popular machine-learning algorithms using super learning to create a prediction function that outperformed the logistic regression model from manuscript 4. This manuscript also found that using a grid search procedure to optimize the hyperparameter values of machine-learning algorithms improved the performance of super learning.

9.2 Main contributions

9.2.1 Substantive contributions

Manuscripts 1 and 2 were the first studies to use data from an indication-based electronic prescribing system to perform a detailed analysis of treatment indications for antidepressants, including off-label indications. The finding that antidepressants were being commonly and increasingly prescribed for indications besides depression that were often not licensed and lacking scientific evidence was eye-opening to physicians, researchers, and the public. These findings were published in high-impact journals (Manuscript 1: *JAMA* in May 2016; Manuscript 2: *The BMJ* in February 2017) where they received an enormous amount of media attention. I was contacted for interviews by over 20 reporters from major local and international news outlets, including TIME magazine (193), Huffington Post (194), CBS News (195,196), Radio Canada International (197), CBC news (198), CTV news (199), and the Montreal Gazette (200). I received emails from researchers at IMS Health in France and Weill Cornell Medicine in New York who were fascinated by the results in manuscript 1 and wanted to reproduce the analysis in their settings. In fact, the researchers from IMS Health have already completed their analysis (which revealed similar findings) and presented the results at the 19th International Society for Pharmacoeconomics and Outcomes Research (ISPOR) European Congress last year. Manuscript 2 had an editorial written on it (also published in *The BMJ*) and was featured on the Figure 1 forum (https://figure1.com) in February 2017, where I did an online Q&A session with health care professionals from around the world. The discussion thread from the online session has since been viewed by over 55,000 physicians and nurses. Finally, I was also contacted by a

physician who was hired by Reuters Health to write a news story for physicians on the findings from manuscript 2. Overall, these experiences demonstrate how the findings from the early papers of this dissertation have raised awareness about off-label antidepressant use, generated discussion about this issue in the medical community, and initiated research on this topic amongst researchers in different countries.

In manuscript 4, I identified several important predictors of antidepressant prescriptions for indications besides depression. These findings may help policy makers and health care managers identify policies and interventions with a higher likelihood of success in changing prescribing behaviours for antidepressants. For example, the finding that the molecule name was by far the strongest predictor of the medical reason why antidepressants were prescribed suggests that policy changes to patient co-pays, formulary decisions, and conditional listings for drugs may have large impacts on prescribing behaviors for antidepressants.

9.2.2 Methodological contributions

Manuscript 3 was the first study to validate the use of diagnostic codes from administrative data to determine antidepressant treatment indications. The classification parameters from this manuscript will provide quantitative estimates to inform bias analyses in studies where diagnostic codes have been used to measure antidepressant treatment indications. Manuscript 3 also demonstrated two valuable considerations when conducting validation studies: 1) the importance of reporting PPV/NPV estimates stratified by factors affecting disease prevalence in the study population, and 2) the benefits of calculating likelihood ratios, which are rarely reported in validation studies but useful for assessing the utility of a test because they are less sensitive than the PPV/NPV to changes in disease prevalence.

Manuscripts 4 and 5 made contributions to improving predictive modelling practices in epidemiology. Manuscript 4 demonstrated a nonconventional model-building procedure that was driven by changes in the scaled Brier score rather than by *p*-values. The benefits of this procedure were that its decisions were based on a metric that directly reflected the predictive

performance of the model and circumvented many of the limitations associated with using *p*-values for model selection, particularly when working with large datasets and testing many covariates (201). Manuscript 5 showed that the performance of the increasingly popular "super learning" prediction methodology could be further improved by optimizing the hyperparameter values of machine-learning algorithms, which does not appear to have been frequently done in the literature.

Finally, this thesis demonstrated a feasible approach for improving the capacity to perform important pharmacovigilance activities for multiple indication drugs that require information on treatment indications. This approach involved identifying a gold-standard source of treatment indications and then deriving algorithms that researchers could use to predict these gold-standard treatment indications using health services data. Manuscripts 4 and 5 showed that it was indeed possible to use health services data to accurately predict when antidepressant prescriptions were written for indications besides depression. I believe that researchers may be able to use the algorithms derived in this thesis to predict treatment indications for antidepressants in settings where treatment indication data is not available.

9.3 Considerations

The key strength of this thesis was the unique data that it used. The availability of gold-standard treatment indications for prescriptions in the MOXXI system linked to a wealth of administrative health data created an unprecedented opportunity to address knowledge gaps and measurement challenges around treatment indications for antidepressants. Besides the unique qualities of the MOXXI data, there are several considerations of this thesis that deserve mention.

9.3.1 New versus prevalent antidepressant users

Pharmacoepidemiology studies often separate new from prevalent drug users (202). In this dissertation, however, I did not distinguish between new and prevalent antidepressant users because drug dispensing data was only available for patients registered in the public drug

insurance plan. For the nearly half (48%) of patients with private drug insurance for whom dispensing data was not available, I did not consider using multiple imputation to impute their therapy status because the characteristics of patients with public drug insurance (e.g. 65+ years old, welfare recipients, unemployed) were notably different from those with private drug insurance (e.g. middle-aged, employed). I also did not restrict the analysis to patients with public drug insurance (for whom therapy status could be determined) because this restriction would have excluded nearly half of the patients in the cohort, thus reducing the generalizability of the study findings. As a result, the findings for new and prevalent antidepressant users separately may differ slightly from the "overall" findings reported in this thesis.

9.3.2 Generalizability of findings

The findings from this thesis were based on prescribing data and thus apply to antidepressant prescriptions. These findings may not be generalizable to *claims* for dispensed antidepressants if the patterns of primary adherence for antidepressants (i.e. filling the prescription) differ by treatment indication. Also, because the MOXXI system is used by primary care physicians in Quebec, the prescribing patterns and diagnostic coding practices of MOXXI physicians may not be generalizable to those of specialists and physicians in other countries. In fact, studies have found that compared to primary care physicians, psychiatrists and other nonpsychiatric physicians are *more* likely to prescribe antidepressants for indications besides depression (14,70). An earlier study (104) also found that half of primary care physicians in the US reported deliberately substituting another diagnostic code for major depression, with the most common reasons being to avoid problems with reimbursement for services or jeopardizing the patient's future ability to obtain health insurance – both of which are not concerns facing physicians in Canada.

9.3.3 Other barriers to performing pharmacovigilance activities for antidepressants

Finally, to conduct an evaluation of antidepressant use for different indications, one must be able to 1) measure the indications for which patients are taking their antidepressants, and 2) measure the outcomes of interest that reflect the safety and effectiveness of the drug. This

thesis attempted to address barriers for measuring treatment indications for antidepressants. However, challenges also exist for measuring the relevant safety and effectiveness outcomes for antidepressants. Adverse drug events are difficult to measure because they are grossly underreported (203) and no universally accepted set of diagnostic codes exists to consistently identify them from administrative data (204). Effectiveness outcomes for common off-label indications for antidepressants such as anxiety disorders, pain, migraine, insomnia, are also difficult – if not impossible – to measure using administrative data. However, the use of natural language processing (NLP) to search for these outcomes in the rich clinical text of physician and nursing notes represents a promising approach for tackling these challenges. In fact, researchers have already begun to explore the use of NLP on clinical text for identifying adverse drugs events (205,206), clinical outcomes in mental health disorders (207), cancer treatment outcomes (208), and even treatment indications for prescriptions (29). NLP is therefore another key methodological approach that should be considered to further enable pharmacovigilance activities for antidepressants.

9.4 Final conclusions and directions for future research

This thesis provides important evidence showing the need for heightened pharmacovigilance around antidepressant use for off-label indications and increases capacity to perform these pharmacovigilance activities by addressing measurement challenges around treatment indications for antidepressants. By building algorithms that use health services data to accurately predict when antidepressants are prescribed for indications besides depression, these algorithms may enable researchers to incorporate information about treatment indications in their analyses, even in the absence of documented treatment indications for antidepressants. Future research work should focus on externally validating the findings from this thesis, particularly the performance of the logistic regression model derived in manuscript 4. To increase the level of precision for predicting specific antidepressant treatment indications, it may also be useful to derive algorithms that predict each treatment indication category separately (e.g. anxiety disorders vs. other, pain vs. other, insomnia vs. other). Finally, to further increase the capacity to conduct comprehensive pharmacovigilance activities for

antidepressants, including risk-benefit evaluations of off-label antidepressant use, research should be done to determine whether NLP can be used on unstructured clinical text to measure safety and effectiveness outcomes for antidepressants.

Appendix A. ICD-9 codes for plausible antidepressant treatment indications

Table A-1. ICD-9 codes for plausible antidepressant treatment indications

Treatment indication	ICD-9 code	Code description	No. prescriptions with diagnostic code recorded[a]	
			Within -3 to +3 days	Within -365 to +3 days
Depressive disorders	296.2	Major depressive disorder, single episode	250	965
	296.3	Major depressive disorder, recurrent episode	169	576
	296.9	Other and unspecified episodic mood disorder	24	462
	300.4	Dysthymic disorder	3467	7886
	301.1	Affective personality disorder (includes chronic depressive personality disorder)	79	258
	309.0	Adjustment disorder with depressed mood	2170	5475
	309.1	Prolonged depressive reaction	60	115
	311.x	Depressive disorder, not elsewhere classified	8317	16265
Anxiety/stress disorders	300.0	Anxiety states	10779	25933
	300.2	Phobic disorders	445	945
	308.x	Acute reaction to stress	130	477
	309.2	Adjustment reaction with predominant disturbance of other emotions	57	379
	309.8	Other specified adjustment reactions (includes posttraumatic stress disorder)	203	434
Sleeping disorders	307.4	Specific disorders of sleep of nonorganic origin	21	123
	780.5	Sleep disturbances	699	3914
Pain	053.1	Herpes zoster with other nervous system complications	11	31
	250.6	Diabetes with neurological manifestations	5	125
	307.8	Pain disorders related to psychological factors	24	202
	337.2	Reflex sympathetic dystrophy	0	0
	338.x	Pain, not elsewhere classified	0	0
	350.1	Trigeminal neuralgia	3	55
	350.2	Atypical face pain	14	53
	352.1	Glossopharyngeal neuralgia	0	0
	353.X	Nerve root and plexus disorders	7	98
	354.x	Mononeuritis of upper limb and mononeuritis multiplex	80	1393
	355.x	Mononeuritis of lower limb	53	603
	357.2	Polyneuropathy in diabetes	34	220

[a]Among all antidepressant prescriptions from the MOXXI system written between January 1, 2003 and December 31, 2012 that were included in the analysis for manuscript 3 (n=77,700).

Table A-1. (continued) ICD-9 codes for plausible antidepressant treatment indications

Treatment Indication	ICD-9 code	Code description	No. prescriptions with diagnostic code recorded[a]	
			Within -3 to +3 days	Within -365 to +3 days
Pain (continued)	714.x	Rheumatoid arthritis and other inflammatory polyarthropathies	250	1113
	715.x	Osteoarthrosis and allied disorders	881	6168
	719.4	Pain in joint	36	452
	721.x	Spondylosis and allied disorders	53	519
	722.x	Intervertebral disc disorders	169	1488
	723.x	Other disorders of cervical region	207	2039
	724.x	Other and unspecified disorders of back	1069	7593
	729.2	Neuralgia, neuritis, and radiculitis, unspecified	86	943
	729.5	Pain in limb	324	4735
	737.x	Curvature of spine	3	44
	786.5	Chest pain	329	5943
	789.0	Abdominal pain	538	7923
Migraine	346.x	Migraine	492	2230
	784.0	Headache	248	2714
Fibromyalgia	729.1	Myalgia and myositis, unspecified	796	2977
Obsessive-compulsive disorder	300.3	Obsessive-compulsive disorder	181	449
Vasomotor symptoms of menopause	627.2	Symptomatic menopausal or female climacteric states	613	3393
Nicotine dependence	305.1	Tobacco use disorder	108	574
Attention deficit/hyperactivity disorder	314.x	Hyperkinetic syndrome of childhood	119	460
Sexual dysfunction	302.7	Psychosexual dysfunction	10	79
	607.8	Other specified disorders of penis	0	27
Pre-menstrual dysphoric disorder	625.4	Premenstrual tension syndromes	26	107
Eating disorders	307.1	Anorexia nervosa	5	44
	307.5	Other and unspecified disorders of eating	8	56
	783.0	Anorexia	15	77
	783.3	Feeding difficulties and mismanagement	4	13

[a]Among all antidepressant prescriptions from the MOXXI system written between January 1, 2003 and December 31, 2012 that were included in the analysis for manuscript 3 (n=77,700).

Appendix B. Supplementary material for manuscript 4

The tables in this Appendix show the 15 most common diagnostic codes, billing code categories, hospital procedure codes, and drugs prescribed in the past year for MOXXI patients that were prescribed antidepressants in manuscript 4.

Table B-1. Top 15 four-digit ICD-9 codes recorded in billings or hospital discharge summary data over the past year for MOXXI patients who were prescribed antidepressants[a]

Four-digit ICD-9 code	Description	Diagnostic code in the past 365 days	
		No. prescriptions	%
V999	No diagnosis or unspecified	25509	34.67
V700	General medical examination	13476	18.32
401.9	Hypertension, unspecified	11309	15.37
401.1	Hypertension, benign	5489	7.46
465.9	Upper respiratory infection, acute, NOSS	4651	6.32
786.0	Dyspnea and respiratory abnormalities	4403	5.98
V723	Special investigations and examinations	3875	5.27
733.0	Osteoporosis	3793	5.16
309.9	Unspecified adjustment reaction	3572	4.85
414.9	Ischemic heart disease	3539	4.81
780.7	Malaise and fatigue	3470	4.72
473.9	Sinusitis, chronic, NOS	3304	4.49
726.9	Unspecified enthesopathy	3223	4.38
692.9	Contact dermatitis, NOS	2915	3.96
300.9	Neurosis, NOS	2680	3.64

[a]After excluding diagnostic codes for plausible antidepressant treatment indications and diagnostic codes for conditions in the Charlson comorbidity index

Table B-2. Top 15 billing code categories from medical billings in the past year for MOXXI patients who were prescribed antidepressants

Description of medical service performed[a]	Category no.	Billing code for the medical service in the past year	
		No. prescriptions	%
Examen complet cabinet - omnipraticien	114	58 323	79.27
Forfaits de responsabilité clientèle vulnérables	170	46 726	63.51
Examen sommaire cabinet - omnipraticien	112	40 449	54.98
Traitements psychiatriques - cabinet	373	34 912	47.45
Examen complet majeur - cabinet	116	24 838	33.76
Actes diagnostiques	211	24 542	33.36
Consultations – spécialistes - établissement	167	23 586	32.06
Examen complet cabinet - spécialiste	115	19 961	27.13
Examen complet – établissement malade externe	141	19 636	26.69
Radiologie	350	19 092	25.95
Consultations – spécialistes – cabinet	165	16 830	22.87
Examen complet – établissement malade externe	135	16 625	22.60
Forfait de prise en charge et de suivi (GMF)	183	16 318	22.18
Examen ordinaire – établissement malade externe	139	15 801	21.48
Examen ordinaire – établissement malade externe	133	15 653	21.27

[a]These are the (French) definitions of the billing code categories that were provided by RAMQ. The descriptions provided by RAMQ were not unique for each billing code group. The category numbers for each billing code category (column 2) were also provided by RAMQ.

Table B-3. Top 15 in-hospital procedures (CCP codes) performed in the past year for MOXXI patients who were prescribed antidepressants

Three-digit CCP code	Description	CCP code in the past year	
		No. prescriptions	%
27.7	Insertion of prosthetic lens	1176	1.60
27.6	Other extraction of lens	1175	1.60
13.0	Transfusion of blood and blood components	632	0.86
13.5	Other injection or infusion of other therapeutic or prophylactic substance	572	0.78
53.0	Bone marrow transplant	536	0.73
48.9	Other operations on vessels of heart	459	0.62
13.6	Respiratory therapy	427	0.58
92.0	Arthrotomy for removal of prosthesis	362	0.49
17.3	Freeing of adhesions and decompression of cranial and peripheral nerves	325	0.44
13.4	Injection or infusion of other therapeutic or prophylactic substance	309	0.42
66.6	Other repair of abdominal wall and peritoneum	266	0.36
90.5	Internal fixation of bone (without fracture reduction)	263	0.36
11.6	Nonoperative removal of therapeutic device from urinary system	258	0.35
93.4	Arthroplasty of knee and ankle	246	0.33
63.1	Cholecystectomy	245	0.33

Abbreviations: CCP = Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP)

Table B-4. Top 15 drugs prescribed in the past year to MOXXI patients who were prescribed antidepressants

Molecule name	Drug prescribed in the past year	
	No. prescriptions	%
Venlafaxine	12674	17.23
Citalopram	10450	14.20
Levothyroxine	7415	10.08
Trazodone	7175	9.75
Atorvastatin	6597	8.97
Lorazepam	6595	8.96
Clonazepam	6382	8.67
Acetaminophen	6186	8.41
Bupropion	6112	8.31
Pantoprazole	5733	7.79
Esomeprazole	5249	7.13
Paroxetine	5205	7.07
Salbutamol	4995	6.79
Amitriptyline	4910	6.67
Metformin	4842	6.58

Appendix C: Copy of published articles included in the thesis

Letters

RESEARCH LETTER

Treatment Indications for Antidepressants Prescribed in Primary Care in Quebec, Canada, 2006-2015

Antidepressant use in the United States has increased over the last 2 decades.[1] A suspected reason for this trend is that primary care physicians are increasingly prescribing antidepressants for nondepressive indications, including unapproved (off-label) indications that have not been evaluated by regulatory agencies.[2] However, the frequency with which physicians prescribe antidepressants for nondepressive indications is unknown because treatment indications are rarely documented. We analyzed the prevalence of treatment indications for antidepressants and assessed temporal trends in antidepressant prescribing for depression.

Methods | This study used data from the Medical Office of the 21st Century (MOXXI) research platform.[3] MOXXI is an electronic medical record (EMR) and prescribing system that has been used by primary care physicians in community-based, fee-for-service practices around 2 major urban centers in Quebec, Canada. During the study period, approximately 185 physicians (25% of eligible) and 100 000 patients (30% of all who visited a MOXXI physician) gave informed consent to use the EMR and have their information used for research purposes.

Table. Treatment Indications and Off-Label Prescribing for Antidepressant Prescriptions in Quebec, Canada, 2006-2015

Treatment Indication[a]	No. of Prescriptions (%)[b]	No. of Prescriptions by Pharmacological Class (%)[c]				No. of Off-Label Prescriptions (%)[h]
		SSRI[d]	SNRI[e]	TCA[f]	Other[g]	
Depressive disorders	56 154 (55.2)	26 339 (46.9)	15 259 (27.2)	1502 (2.7)	13 054 (23.3)	0 (0.0)
Anxiety disorders[i]	18 849 (18.5)	12 466 (66.1)	5076 (26.9)	273 (1.5)	1034 (5.5)	8975 (47.6)
Insomnia	10 347 (10.2)	19 (0.2)	2 (0.0)	2242 (21.7)	8804 (78.1)	10 077 (97.4)
Pain	6241 (6.1)	24 (0.4)	1340 (21.5)	4623 (74.1)	254 (4.1)	5174 (82.9)
Panic disorders with or without agoraphobia	4174 (4.1)	3280 (78.6)	751 (18.0)	89 (2.1)	54 (1.3)	1487 (35.6)
Fibromyalgia	1550 (1.5)	63 (4.1)	958 (61.8)	506 (32.7)	23 (1.5)	946 (61.0)
Migraine	1498 (1.5)	6 (0.4)	22 (1.5)	1470 (98.1)	0 (0.0)	1498 (100.0)
Obsessive-compulsive disorder	1111 (1.1)	875 (78.8)	177 (15.9)	53 (4.8)	6 (0.5)	435 (39.2)
Vasomotor symptoms of menopause	856 (0.8)	112 (13.1)	736 (86.0)	2 (0.2)	6 (0.7)	856 (100.0)
Social phobia	568 (0.6)	434 (76.4)	134 (23.6)	0 (0.0)	0 (0.0)	199 (35.0)
Nicotine dependence	514 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)	514 (100.0)	0 (0.0)
Attention-deficit/hyperactivity disorder	389 (0.4)	16 (4.1)	4 (1.0)	5 (1.3)	364 (93.6)	389 (100.0)
Posttraumatic stress disorder	263 (0.3)	211 (80.2)	35 (13.3)	2 (0.8)	15 (5.7)	207 (78.7)
Sexual dysfunction	261 (0.3)	39 (14.9)	2 (0.8)	8 (3.1)	212 (81.2)	261 (100.0)
Premenstrual disorders and syndromes	212 (0.2)	193 (91.0)	17 (8.0)	2 (0.9)	0 (0.0)	63 (29.7)
Digestive system disorders	119 (0.1)	4 (3.4)	1 (0.8)	114 (95.8)	0 (0.0)	119 (100.0)
Urinary system disorders	109 (0.1)	0 (0.0)	2 (1.8)	107 (98.2)	0 (0.0)	100 (91.7)
Bulimia nervosa	76 (0.1)	54 (71.1)	0 (0.0)	12 (15.8)	10 (13.2)	22 (29.0)
Other	317 (0.3)	61 (19.2)	11 (3.5)	126 (39.8)	119 (37.5)	211 (66.6)
Any indication	101 759 (100.0)	43 462 (42.7)	23 898 (23.5)	10 936 (10.8)	23 463 (23.1)	29 907 (29.4)

Abbreviations: SNRI, serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

[a] Among all antidepressant prescriptions, 1.8% had multiple treatment indications recorded and were assigned to multiple categories. As a result, the sum of prescriptions across the individual treatment indication categories exceeds the number of prescriptions for any indication (last row).

[b] Percentages were calculated using the total number of antidepressant prescriptions for any indication (N = 101 759) as the denominator.

[c] Percentages for each pharmacological class were calculated using the total number of prescriptions for the indication as the denominator.

[d] Includes citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline.

[e] Includes desvenlafaxine, duloxetine, and venlafaxine.

[f] Includes amitriptyline, clomipramine, desipramine, doxepin, imipramine, nortriptyline, and trimipramine.

[g] Includes bupropion, maprotiline, mirtazapine, trazodone, and vortioxetine.

[h] For each treatment indication category, a prescription was classified as off-label if the drug was not approved for the indication by Health Canada or the US Food and Drug Administration as of September 2015. For any indication (last row), a prescription was classified as off-label if the drug was not approved for all of its recorded indications. Percentages were calculated using the total number of prescriptions for the indication as the denominator.

[i] Includes anxiety, generalized anxiety disorder, and other anxiety disorders except panic disorders and phobias.

Compared with nonconsenters, MOXXI physicians were younger and MOXXI patients were older with more health complexities.[4]

This study included all prescriptions written for adults between January 1, 2006, and September 30, 2015, for all antidepressants except monoamine oxidase inhibitors. Physicians had to document at least 1 treatment indication per prescription using a drop-down menu containing a list of indications or by typing the indication(s). In a validation study, these indications had excellent sensitivity (98.5%) and high positive predictive value (97.0%).[5] Prescriptions were classified as on-label or off-label depending on whether the drug was approved for the indication by Health Canada or the US Food and Drug Administration by September 2015. Temporal trends in antidepressant prescribing for depression were measured using generalized linear risk difference models for binary outcomes, with an identity link. A linear effect of calendar time (in years) was modeled on the probability of antidepressant prescribing for depression, adjusted for patient age and sex and accounting for multilevel clustering of prescriptions using an alternating logistic regression algorithm.[6] All statistical analyses were conducted using SAS (SAS Institute) software, version 9.4. This study was approved by the McGill institutional review board.

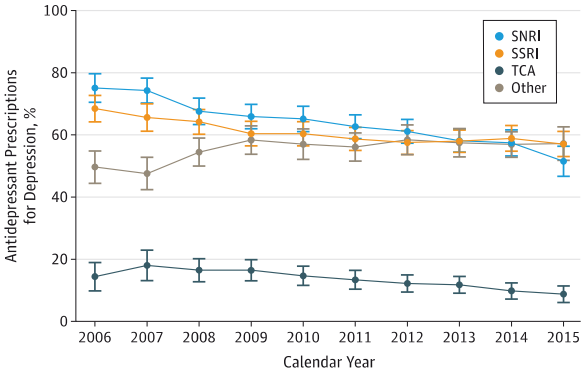
Results | During the study period, 101 759 antidepressant prescriptions (5.9% of all prescriptions) were written by 158 physicians for 19 734 patients. Only 55.2% of antidepressant prescriptions were indicated for depression. Physicians also prescribed antidepressants for anxiety disorders (18.5%), insomnia (10.2%), pain (6.1%) and panic disorders (4.1%) (**Table**). For these indications, respectively, the most frequently prescribed antidepressants were citalopram (29.5% of prescriptions for the indication), trazodone (76.6%), amitriptyline (65.1%), and paroxetine (35.9%).

For 29.4% of all antidepressant prescriptions (65.6% of prescriptions not for depression), physicians prescribed a drug for an off-label indication, especially insomnia and pain. Physicians also prescribed antidepressants for several indications that were off-label for all antidepressants, including migraine, vasomotor symptoms of menopause, attention-deficit/hyperactivity disorder, and digestive system disorders (Table).

Between 2006 and 2015, the percentage of antidepressants prescribed for depression decreased significantly, with an adjusted 5-year risk difference of −9.73% (95% CI, −11.86% to −7.61%) for serotonin-norepinephrine reuptake inhibitors, −3.96% (95% CI, −5.33% to −2.59%) for selective serotonin reuptake inhibitors, and −2.99% (95% CI, −4.90% to −1.08%) for tricyclic antidepressants (**Figure**). However, the percentage of other antidepressants (especially mirtazapine) prescribed for depression increased significantly, with an adjusted 5-year risk difference of 2.36% (95% CI, 0.32% to 4.40%).

Discussion | Between 2006 and 2015, primary care physicians in Quebec commonly and increasingly prescribed antidepressants for nondepressive indications. When physicians prescribed antidepressants for insomnia and pain, they often prescribed antidepressants off-label.

Figure. Percentage of Antidepressant Prescriptions for Depression by Pharmaceutical Class, 2006-2015



Total No. of antidepressant prescriptions

2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
7541	9216	9893	11315	11964	12121	11610	11313	9974	6812

SNRI indicates serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

The plots show the unadjusted percentage of antidepressant prescriptions written for depression in each calendar year by pharmacological class. The error bars represent 95% CIs that were calculated based on standard errors corrected for multilevel clustering of prescriptions using an alternating logistic regression algorithm.[6] Five-year risk difference estimates in the percentage of antidepressant prescriptions for depression were obtained from a generalized linear risk difference model (with an identity link) that included a linear effect of calendar time and dummy variables for individual pharmaceutical classes, along with their interactions with calendar time. All risk-difference estimates were adjusted for patient age and sex and used an alternating logistic regression algorithm[6] to account for multilevel clustering of prescriptions. As there were no missing data on patient age or sex, all prescriptions were included in the regression model.

The study was limited by a selective patient population and a small number of prescribers from 1 Canadian province. However, this is the first study to our knowledge to describe the prevalence of treatment indications for antidepressants using validated, physician-documented treatment indications recorded at the point of prescribing. The findings indicate that the mere presence of an antidepressant prescription is a poor proxy for depression treatment, and they highlight the need to evaluate the evidence supporting off-label antidepressant use.

Jenna Wong, MSc
Aude Motulsky, PhD
Tewodros Eguale, MD, PhD
David L. Buckeridge, MD, PhD
Michal Abrahamowicz, PhD
Robyn Tamblyn, PhD

Author Affiliations: Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada (Wong, Motulsky, Buckeridge, Abrahamowicz, Tamblyn); School of Pharmacy, Massachusetts College of Pharmacy and Health Sciences, Boston (Eguale).

Corresponding Author: Jenna Wong, MSc, Faculty of Medicine, Clinical and Health Informatics Research Group, Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, 1140 Pine Ave W, Montreal, Quebec, H3A 1A3 (jenna.wong@mail.mcgill.ca).

Author Contributions: Dr Tamblyn and Ms Wong had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Wong, Motulsky, Tamblyn.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Wong.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Wong, Abrahamowicz.
Obtained funding: Wong, Tamblyn.
Administrative, technical, or material support: Tamblyn.
Study supervision: Tamblyn.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Role of the Funder/Sponsor: The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript, and decision to submit the manuscript for publication.

1. National Center for Health Statistics. Health, United States, 2010 with special feature on death and dying. http://www.cdc.gov/nchs/data/hus/hus10.pdf. Accessed July 23, 2015.

2. Gardarsdottir H, Heerdink ER, van Dijk L, Egberts ACG. Indications for antidepressant drug prescribing in general practice in the Netherlands. *J Affect Disord*. 2007;98(1-2):109-115.

3. 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.

4. Bartlett G, Tamblyn R, Kawasumi Y, Poissant L, Taylor L. Nonparticipation bias in health services research using data from an integrated electronic prescribing project: the role of informed consent. *Acta Bioeth*. 2005;11(2):145-159. http://www.scielo.cl/pdf/abioeth/v11n2/art05.pdf.

5. 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.

6. Carey V, Zeger SL, Diggle P. Modelling multivariate binary data with alternating logistic regressions. *Biometrika*. 1993;80(3):517-526. doi:10.2307/2337173.

COMMENT & RESPONSE

Risk of Anaphylaxis With Intravenous Iron Products

To the Editor The study by Dr Wang and colleagues[1] reported the risk of anaphylaxis with different intravenous iron products. I have several concerns.

One concern is the definition of *anaphylaxis*. In the article's online Supplement, both anaphylaxis criteria B and C include "injection of diphenhydramine." Given that diphenhydramine often is given both as a premedication (without evidence to support its use) and for treatment of minor reactions (flushing, myalgias of the chest and back, or headache), including it in the definition has the potential to greatly overestimate anaphylaxis rates. Diphenhydramine itself has been reported to cause somnolence, tachycardia, and hypotension.[2] Published data suggest that when diphenhydramine is used as premedication prior to intravenous iron, a significant majority of the reactions ostensibly attributed to the iron were due to diphenhydramine.[3] In the Methods section, the authors noted that minor self-limited reactions such as minor allergic reactions would be wrongly labeled as anaphylaxis events. It would be informative to know the rates of anaphylactic reactions when patients who received diphenhydramine were excluded.

A second concern is that these data need to be put into perspective by comparing intravenous iron products with other commonly used drugs. For example, anaphylaxis to penicillin is reported in 10 to 50 per 100 000 treatment courses with a fatality rate of 1 to 2 per 100 000 courses, which is in the same range as that reported for intravenous iron.[4]

Third, death rates, which were reported in the article's online Supplement, were not discussed in the article. Reviewing data from eTables 3 and 4, it appears that death rates for all iron products were very similar, suggesting that safety rates are likely to be similar for all products and that the anaphylaxis rates for iron dextran may have been overestimated. Is it possible that criteria by which anaphylaxis was defined differed for different products?

Finally, it should be emphasized that the study population, with an average age of 73 to 74 years, was not representative of many patients receiving intravenous iron, including patients who have inflammatory bowel disease or heavy uterine bleeding, are pregnant, or are undergoing bariatric surgery. It is important not to extrapolate rates from older adults to these younger (and potentially healthier) individuals.

Thomas G. DeLoughery, MD, MACP, FAWM

Author Affiliation: Division of Hematology/Medical Oncology, Oregon Health Sciences University, Portland.

Corresponding Author: Thomas G. DeLoughery, MD, MACP, FAWN, Division of Hematology/Medical Oncology, Oregon Health Sciences University, 3181 SW Sam Jackson Park Road, Hematology L586, Portland, OR 97201-3098 (deloughT@ohsu.edu).

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

1. Wang C, Graham DJ, Kane RC, et al. Comparative risk of anaphylactic reactions associated with intravenous iron products. *JAMA*. 2015;314(19):2062-2068.

2. Agostini JV, Leo-Summers LS, Inouye SK. Cognitive and other adverse effects of diphenhydramine use in hospitalized older patients. *Arch Intern Med*. 2001;161(17):2091-2097.

3. Barton JC, Barton EH, Bertoli LF, Gothard CH, Sherrer JS. Intravenous iron dextran therapy in patients with iron deficiency and normal renal function who failed to respond to or did not tolerate oral iron supplementation. *Am J Med*. 2000;109(1):27-32.

4. Lieberman P, Camargo CA Jr, Bohlke K, et al. Epidemiology of anaphylaxis: findings of the American College of Allergy, Asthma and Immunology Epidemiology of Anaphylaxis Working Group. *Ann Allergy Asthma Immunol*. 2006;97(5):596-602.

In Reply In response to Dr DeLoughery's concern about diphenhydramine, provided is a sensitivity analysis that excluded all anaphylaxis cases with the claim of "injection of diphenhydramine" (**Table 1**). Results are consistent with the primary analysis.

Regarding incidence rates, as discussed in our article, rates from our study were substantially lower than those reported from clinical trials of intravenous iron products. Although other study populations might have contributed to the observed differences, other differences in identifying and reporting anaphylaxis cases during clinical trials and during general clinical practice might also be relevant. Practicing physicians might not classify less severe or atypical anaphy-

Off-label indications for antidepressants in primary care: descriptive study of prescriptions from an indication based electronic prescribing system

Jenna Wong,[1] Aude Motulsky,[1,2] Michal Abrahamowicz,[1] Tewodros Eguale,[3] David L Buckeridge,[1] Robyn Tamblyn[1]

[1]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montréal, Canada
[2]Centre de recherche du Centre hospitalier de l'Université de Montréal, School of Public Health, University of Montréal, Montréal, Canada
[3]Massachusetts College of Pharmacy and Health Sciences University, Boston, MA, USA
Correspondence to: J Wong
jenna.wong@mail.mcgill.ca

Cite this as: *BMJ* 2017;356:j603
http://dx.doi.org/10.1136/bmj.j603
Accepted: 18 January 2017

ABSTRACT

OBJECTIVE
To examine off-label indications for antidepressants in primary care and determine the level of scientific support for off-label prescribing.

DESIGN
Descriptive study of antidepressant prescriptions written by primary care physicians using an indication based electronic prescribing system.

SETTING
Primary care practices in and around two major urban centres in Quebec, Canada.

PARTICIPANTS
Patients aged 18 years or older who visited a study physician between 1 January 2003 and 30 September 2015 and were prescribed an antidepressant through the electronic prescribing system.

MAIN OUTCOME MEASURES
Prevalence of off-label indications for antidepressant prescriptions by class and by individual drug. Among off-label antidepressant prescriptions, the proportion of prescriptions in each of the following categories was measured: strong evidence supporting use of the prescribed drug for the respective indication; no strong evidence for the prescribed drug but strong evidence supporting use of another drug in the same class for the indication; or no strong evidence supporting use of the prescribed drug and all other drugs in the same class for the indication.

RESULTS
106 850 antidepressant prescriptions were written by 174 physicians for 20 920 adults. By class, tricyclic antidepressants had the highest prevalence of off-label indications (81.4%, 95% confidence interval, 77.3% to 85.5%), largely due to a high off-label prescribing rate for amitriptyline (93%, 89.6% to 95.7%). Trazodone use for insomnia was the most common off-label use for antidepressants, accounting for 26.2% (21.9% to 30.4%) of all off-label prescriptions. For only 15.9% (13.0% to 19.3%) of all off-label prescriptions, the prescribed drug had strong scientific evidence for the respective indication. For 39.6% (35.7% to 43.2%) of off-label prescriptions, the prescribed drug did not have strong evidence but another antidepressant in the same class had strong evidence for the respective indication. For the remaining 44.6% (40.2% to 49.0%) of off-label prescriptions, neither the prescribed drug nor any other drugs in the class had strong evidence for the indication.

CONCLUSIONS
When primary care physicians prescribed antidepressants for off-label indications, these indications were usually not supported by strong scientific evidence, yet often another antidepressant in the same class existed that had strong evidence for the respective indication. There is an important need to generate and provide physicians with evidence on off-label antidepressant use to optimise prescribing decisions.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Off-label drug use without strong scientific evidence is associated with an increased risk of adverse drug events
About a third of all antidepressants in primary care are prescribed for off-label indications
The degree to which off-label antidepressant prescriptions are supported by strong scientific evidence is unknown

WHAT THIS STUDY ADDS

Most off-label antidepressant prescriptions lack strong scientific evidence, but another evidence based antidepressant from the same class could often be considered as an alternative
There is an important need to produce more evidence evaluating the clinical outcomes associated with off-label antidepressant use
Indication based electronic prescribing systems represent an effective means to study off-label antidepressant use and communicate evidence back to physicians to optimise prescribing decisions

Introduction

Antidepressant use has increased substantially in the UK[1,2] and in other western countries such as Canada[3] and the USA.[4] In fact, the number of antidepressants dispensed in England increased by 3.9 million (6.8%) between 2014 and 2015—more than any other therapeutic class of prescription drugs.[2] One suspected factor underlying the widespread use of antidepressants is an expanding array of indications for these drugs, many of which are unapproved (off-label) for certain antidepressants.[5]

There is a lack of epidemiological evidence on the extent to which physicians prescribe antidepressants for off-label indications because treatment indications are not documented for most prescriptions.[6] With the advent of electronic prescribing (e-prescribing) systems, however, formal documentation of treatment indications linked to prescriptions (that is, indication based prescribing) is possible. Although indication based prescribing is not broadly used at the moment, it represents a valuable means for studying off-label prescribing.[7] We recently used data from a unique, indication based e-prescribing system to describe treatment

indications for antidepressants in primary care.[8] We found that over the past decade, primary care physicians commonly and increasingly prescribed antidepressants for non-depressive indications. Moreover, when antidepressants were not prescribed for depression, two of three prescriptions were for an off-label indication.

Off-label prescribing warrants particular attention and oversight when the drug use is not supported by scientific evidence showing greater benefits relative to risk.[9 10] Inefficacious antidepressant use is a concern because it creates unnecessary costs and puts patients at risk of experiencing burdensome side effects and serious adverse events that could be avoided. For example, even though newer generation antidepressants such as selective serotonin reuptake inhibitors (SSRIs) are considered safer and more tolerable than the older generation tricyclic antidepressants (TCAs), they are costly and have still been associated with notable side effects and safety concerns. These side effects include sexual dysfunction, drowsiness, insomnia, weight gain, and fatigue,[11-13] and safety concerns include an increased risk of fractures[14] and upper gastrointestinal bleeds.[15 16] Off-label antidepressant use could also expose patients to unknown health risks if their clinical characteristics differ from the patient populations studied in pre-market clinical trials.[17] Indeed, the risk of adverse drug events has been found to be 54% higher when drugs are used off-label without strong scientific evidence than when drugs are used on-label.[18]

Although an estimated 29% of antidepressants are prescribed for off-label indications,[8] it is unknown to what extent these off-label prescriptions are supported by scientific evidence. Thus, the objective of this study was to examine off-label indications for antidepressants in primary care and assess the level of scientific evidence supporting these off-label prescriptions.

Methods

Study design and setting

This descriptive study took place in the Canadian province of Quebec, where a universal health insurance programme covers the cost of essential medical care for all residents. By law, all residents must be covered for prescription drugs through either private plans (that is, group or employee benefit plans) or the public drug insurance plan. About 50% of residents are registered in the public drug insurance plan, including those older than 65, welfare recipients, and those not insured through an employer. At a minimum, all private plans must provide the same formulary for insured drugs as the public drug insurance plan.[19]

Data source and study population

The Medical Office of the XXIst Century (MOXXI) is an electronic prescription and drug management system used by consenting primary care physicians in community based, fee-for-service practices around two major urban centres in Quebec.[20] Since 2003, 207 physicians (25% of eligible physicians) and over 100 000 patients (26% of all who visited a MOXXI physician) have consented to participate in the MOXXI programme and have their information used for research purposes.

The e-prescribing tool in the MOXXI system requires physicians to explicitly record at least one treatment indication per prescription by either using a dropdown menu that lists on-label and off-label indications (without distinction) or typing the indication(s) into a free text field. In a validation study,[21] these physician documented indications had excellent sensitivity (98.5%) and high positive predictive value (97.0%) when compared with a blinded, post hoc, physician facilitated chart review. The MOXXI system also provides physicians with access to professional drug monographs that are maintained by a commercial vendor[22] and produces automated drug alerts about potential prescribing problems. Alerts are generated when potential dosing errors or drug-drug, drug-disease, drug-age, or drug-allergy contraindications are identified; however, alerts are not generated when drugs are prescribed for off-label indications. This study was approved by the McGill institutional review board.

Inclusion and exclusion criteria

This study included prescriptions of drugs approved for depression that were written by MOXXI physicians between 1 January 2003 and 30 September 2015 for patients aged 18 years or older. The antidepressant prescription was the unit of analysis. We excluded drugs with fewer than 150 prescriptions during the study period (roughly corresponding to a prescribing frequency of fewer than once per month). This resulted in the exclusion of all monoamine oxidase inhibitors (phenelzine, tranylcypromine, moclobemide, and isocarboxazid), nefazodone, maprotiline, and vortioxetine.

Measurements

On-label versus off-label indications

Treatment indications were first categorised by use of ICD-10 (international classification of diseases, 10th revision). Each prescription—representing a drug-indication pair—was then classified as on-label or off-label, depending on whether the drug had been approved for the indication by Health Canada or the US Food and Drug Administration as of September 2015 (the end of the study period). Approved indications were determined at the end of the study period rather than the year in which the prescription was written so that all prescriptions would be classified using the same benchmark. If a physician recorded multiple indications for the drug (n=1922, 1.8% of all antidepressant prescriptions), the prescription was classified as off-label only if all the indications were not approved.

Level of scientific evidence for off-label prescriptions

Off-label prescriptions were further analysed according to the level of scientific evidence supporting the drug's use for the off-label indication. Off-label prescriptions were assigned to one of three categories: strong evidence for the prescribed drug, no strong evidence for the prescribed drug but strong evidence for another drug in the same class, or no strong evidence for the

doi: 10.1136/bmj.j603 | *BMJ* 2017;356:j603 | the bmj

prescribed drug and all other drugs in the same class. To determine whether off-label prescriptions had strong evidence for the prescribed drug, we used the DRUGDEX compendium (Thomson Micromedex),[23] which is a reputable and authoritative reference used by the US Centers for Medicare and Medicaid Services to determine coverage for off-label drug uses.[24] The compendium contains evaluations of drug efficacy, strength of recommendation, and strength of evidence for off-label drug indication pairs.

Using the same criteria as in previous studies,[7 18 25] we classified off-label prescriptions as having strong evidence for the prescribed drug if evidence showed that the drug was effective or favoured efficacy for the indication, the drug was recommended for all or most patients with the indication, and at least one randomised clinical trial was included among the studies used to evaluate the drug's efficacy for the indication. If an off-label prescription did not have strong evidence for the prescribed drug, we then determined whether there was strong evidence for another drug in the same class. This condition was satisfied if another drug in the same class was either on-label or off-label with strong evidence for the indication. If an off-label prescription still did not have strong evidence for another drug in the class, then the prescription was classified as having no strong evidence for the prescribed drug and all other drugs in the same class.

Statistical analysis

Patient and physician characteristics were summarised by use of descriptive statistics. The prevalence of off-label indications was estimated as the number of off-label prescriptions divided by the total number of antidepressant prescriptions overall, in the class, or for the individual drug. We estimated the level of scientific evidence for off-label prescriptions as the number of off-label prescriptions in each evidence category divided by the total number of off-label antidepressant prescriptions overall or in the class. The prevalence of different treatment indications for each drug was estimated as a proportion, using the total number of prescriptions for the drug as the denominator. For all proportions, we calculated 95% confidence intervals using a cluster bootstrap approach[26] to account for within-cluster correlation among prescriptions for the same patient and from the same physician. The reported 95% confidence intervals correspond to the values of the 2.5th and 97.5th percentiles of the distribution of the respective estimates across 1000 bootstrap re-samples.[26] All analyses were conducted using SAS (SAS Institute) software, version 9.4.

Patient involvement

No patients were involved in setting the research question or the study measures, nor were they involved in developing plans for the design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. The study findings will be disseminated to study participants through physician newsletters and patient-friendly handouts.

Results

During the study period, 106 850 antidepressant prescriptions (5.8% of 1.83 million prescriptions for any drug) were written by 174 primary care physicians for 20 920 adults. There was roughly an equal number of male (n=90; 52%) and female (n=84; 48%) physicians, most of whom had been trained in North America (n=160; 92%) and practicing for at least 15 years (n=131; 75%). Two thirds of patients were female (n=13 990; 66.9%), most patients were middle aged at the time of their earliest antidepressant prescription (median 53 years, interquartile range 43-65), and patients were equally likely to have public (n=10 875; 52.0%) or private (n=10 045; 48.0%) drug insurance. Over the study period, patients had a median of three (interquartile range 1-7) antidepressant prescriptions and were prescribed a median of one (1-2) type of antidepressant drug.

Prevalence of off-label indications

Overall, 29.3% (95% confidence interval 26.6% to 32.3%) of all antidepressant prescriptions were written for an off-label indication (table 1). By class, TCAs had the highest prevalence of off-label indications (81.4%, 77.3% to 85.5%), followed by other antidepressants (trazodone, bupropion, and mirtazapine; 42.4%, 37.1% to 47.7%) and SSRIs (21.8%, 19.0% to 25.0%). By contrast, the prevalence of off-label indications was much lower for serotonin-norepinephrine (noradrenaline) reuptake inhibitors (SNRIs; 6.1%, 4.8% to 7.5%). The high prevalence of off-label indications for TCAs was mostly due to amitriptyline, which was only approved for depression but was almost exclusively prescribed for off-label indications (93.0%, 89.6% to 95.7%)—most commonly pain (48.4%, 39.7% to 57.8%), insomnia (22.5%, 13.6% to 31.3%), and migraine (16.7%, 12.2% to 21.9%; table 2). The high prevalence of off-label indications among other antidepressants (trazodone, bupropion, and mirtazapine) was largely due to trazodone, which was mostly prescribed for insomnia (82.5%, 74.5% to 88.1%) even though it was not approved for this indication. SSRIs and SNRIs had a lower prevalence of off-label indications because they were more frequently prescribed for depression than TCAs, which by definition was an approved indication for all antidepressants (table 2).

Level of scientific evidence for off-label indications

Among all off-label antidepressant prescriptions, there were 143 unique drug indication pairs—the most common of which were trazodone for insomnia (representing 26.2%, 95% confidence interval 21.9% to 30.4%, of all off-label prescriptions), citalopram for anxiety (17.8%, 14.8% to 21.3%), amitriptyline for pain (13.8%, 11.0% to 16.9%), and amitriptyline for insomnia (6.4%, 3.9% to 9.5%; data not shown). Only three of these 143 off-label drug indication pairs met the predefined criteria[7 18 25] for having strong scientific evidence: amitriptyline (a TCA) for pain, escitalopram (an SSRI) for panic disorders, and venlafaxine (an SNRI) for obsessive compulsive disorder.

Table 1 | Proportion of antidepressants prescribed for off-label indications and level of evidence, by drug class

Drug class (No of prescriptions)	Off-label indication		Level of evidence for off-label indications					
			Strong evidence for prescribed drug‡		No strong evidence for prescribed drug but strong evidence for another drug in same class¶		No strong evidence for prescribed drug and all other drugs in same class	
	No	Percentage* (95% CI†)	No	Percentage§ (95% CI†)	No	Percentage§ (95% CI†)	No	Percentage§ (95% CI†)
SSRI (n=45 608)	9960	21.8 (19.0 to 25.0)	473	4.7 (2.7 to 7.2)	9160	92.0 (89.2 to 94.4)	327	3.3 (2.0 to 4.8)
SNRI (n=25 235)	1539	6.1 (4.8 to 7.5)	169	11.0 (4.6 to 18.4)	544	35.4 (25.0 to 46.7)	826	53.7 (40.6 to 66.6)
TCA (n=11 645)	9480	81.4 (77.3 to 85.5)	4335	45.7 (37.8 to 54.0)	2682	28.3 (20.5 to 36.6)	2463	26.0 (21.2 to 31.1)
Other** (n=24 362)	10 340	42.4 (37.1 to 47.7)	0	0.0 (0.0 to 0.0)	NA	NA	10 340	100.0 (100.0 to 100.0)
All classes (n=106 850)	31 319	29.3 (26.6 to 32.3)	4977	15.9 (13.0 to 19.3)	12 386	39.6 (35.7 to 43.2)	13 956	44.6 (40.2 to 49.0)

SSRI=selective serotonin reuptake inhibitors; SNRI=serotonin-norepinephrine reuptake inhibitors; TCA=tricyclic antidepressants; NA=not assessed for drugs in this category because they were not considered as part of the same class.
*Calculated using the total number of prescriptions in the class as the denominator.
†Calculated by a cluster bootstrap approach[26] to account for non-independence of prescriptions from the same physician and for the same patient. Reported 95% confidence intervals correspond to values at the 2.5th and 97.5th percentiles of the distribution of respective estimates across 1000 bootstrap re-samples.
‡Based on evaluations from DRUGDEX compendium in three dimensions: efficacy, strength of recommendation, and strength of evidence. Prescriptions for an off-label indication were classified as having strong evidence for a prescribed drug if evidence showed that the drug was effective or favoured efficacy for the indication, the drug was recommended for all or most patients with the indication, and at least one randomised controlled trial was included among the studies used to evaluate the drug's efficacy for the indication.
§Calculated using the number of prescriptions in the class that were written for an off-label indication as the denominator.
¶Off-label prescriptions where the prescribed drug did not have strong evidence for the indication, but another drug in the class was either on-label or off-label with strong evidence for the indication based on evaluations from the DRUGDEX compendium.
**Includes trazodone, bupropion, and mirtazapine.

These three pairs collectively comprised 15.9% (13.0% to 19.3%) of all off-label antidepressant prescriptions (table 1)—most which were amitriptyline prescriptions for pain (representing 87.1%, 80.9% to 92.1%, of all off-label prescriptions with strong evidence for the prescribed drug). As a result, the proportion of off-label antidepressant prescriptions with strong evidence for the prescribed drug was much higher for TCAs (45.7%, 37.8% to 54.0%) than for SNRIs (11.0%, 4.6% to 18.4%) and SSRIs (4.7%, 2.7% to 7.2%; table 1).

Off-label antidepressant prescriptions had strong evidence for another drug in the same class—but not the prescribed drug—in 39.6% (95% confidence interval 35.7% to 43.2%) of all cases (table 1). This proportion was highest among off-label SSRI prescriptions (92.0%, 89.2% to 94.4%), and lower among off-label prescriptions for SNRIs (35.4%, 25.0% to 46.7%) and TCAs (28.3%, 20.5% to 36.6%). This proportion was not assessed for other antidepressants because trazodone, bupropion, and mirtazapine were not considered as part of the same class.

For the remaining 44.6% (95% confidence interval 40.2% to 49.0%) of off-label antidepressant prescriptions, neither the prescribed drug nor any other drug in the same class had strong evidence for the indication (table 1). All off-label prescriptions for other antidepressants (trazodone, bupropion, and mirtazapine) were classified in this evidence category. The proportion of off-label prescriptions with no scientific support for any drug in the class was also quite high for SNRIs (53.7%, 40.6% to 66.6%) and TCAs (26.0%, 21.2% to 31.1%), but was much lower for SSRIs (3.3%, 2.0% to 4.8%).

Discussion

This study provides evidence on the level of scientific support for off-label antidepressants prescriptions, the prevalence of off-label indications for individual antidepressants, and the most common off-label uses for antidepressants. Nearly a third (29%) of all antidepressants in this study were prescribed for an off-label indication, as found previously.[8] Among all off-label antidepressant prescriptions, only one in six prescriptions was supported by strong scientific evidence, but there was often another antidepressant in the same class with strong evidence that could have been considered instead, especially among off-label SSRI prescriptions. Still, nearly half of all off-label antidepressant prescriptions did not have strong evidence for the prescribed drug and all other antidepressants in the same class. Among the many off-label uses for antidepressants, physicians most frequently prescribed trazodone for insomnia even though this use was not evidence based.

Comparison with other studies

Few published studies exist on off-label prescribing, owing to challenges associated with measuring diagnoses (indications) for prescriptions. Compared with our findings where 29% of antidepressant prescriptions were off-label, Chen and colleagues[27] found that 75% of people enrolled to Georgia Medicaid who were being treated with antidepressants received at least one antidepressant off-label. The rate of off-label antidepressant use was notably higher in this study because the authors classified prescriptions as off-label if the patient did not have a diagnostic code for an approved indication recorded in administrative claims data during the same year. This methodology most likely overestimated the off-label prescribing rate because diagnostic codes in administrative data are often incomplete or inaccurate, especially for psychiatric conditions.[28]

Only three studies—one Canadian[7] and two US[25 29]—have used documented treatment indications to study off-label prescribing, none of which focused specifically on antidepressants. Eguale and colleagues[7] combined antidepressants with other central nervous system drugs but reported fairly comparable results, with 26% of prescriptions for off-label indications—18% of which

doi: 10.1136/bmj.j603 | BMJ 2017;356:j603 | the bmj

Table 2 | Off-label indications and most common indications for antidepressant treatment, by drug

Drug name, by class	Total No of prescriptions	Off-label indication		Treatment indications and No (%) of prescriptions								
				Most common			Second most common			Third most common		
		No	Percentage* (95% CI†)	Indication	No	Percentage* (95% CI†)	Indication	No	Percentage* (95% CI†)	Indication	No	Percentage* (95% CI†)
Selective serotonin reuptake inhibitors												
Citalopram	19480	6988	35.9 (31.5 to 40.9)	Depression‡	12492	64.1 (59.1 to 68.5)	Anxiety disorder§	5745	29.5 (25.0 to 34.6)	Panic disorder	882	4.5 (2.7 to 7.2)
Paroxetine	9212	94	1.0 (0.4 to 1.9)	Depression‡	4476	48.6 (40.2 to 57.3)	Anxiety disorder‡§	2719	29.5 (23.6 to 36.0)	Panic disorder‡	1563	17.0 (10.6 to 23.8)
Escitalopram	7108	601	8.5 (5.7 to 11.4)	Depression‡	4354	61.3 (55.2 to 67.0)	Anxiety disorder‡§	2075	29.2 (23.2 to 35.3)	Panic disorder	503	7.1 (4.6 to 9.9)
Sertraline	6805	1680	24.7 (18.6 to 31.8)	Depression‡	4383	64.4 (55.4 to 71.8)	Anxiety disorder§	1847	27.1 (20.6 to 34.1)	Panic disorder‡	398	5.8 (2.4 to 10.0)
Fluoxetine	2079	322	16.0 (10.0 to 24.5)	Depression‡	1566	75.3 (65.1 to 83.2)	Anxiety disorder§	249	12.0(7.1 to 18.2)	Panic disorder‡	64	3.1 (0.2 to 7.4)
Fluvoxamine	924	265	28.7 (15.6 to 45.3)	Depression‡	592	64.1 (47.5 to 76.5)	Anxiety disorder§	233	25.2 (13.8 to 41.9)	OCD‡	71	7.7 (2.9 to 15.5)
Serotonin-norepinephrine reuptake inhibitors												
Venlafaxine	21369	1501	7.0 (5.5 to 8.7)	Depression‡	14282	66.8 (62.3 to 71.2)	Anxiety disorder‡§	5053	23.6 (19.4 to 27.9)	Panic disorder‡	782	3.7 (2.6 to 4.9)
Duloxetine	2969	9	0.3 (0.0 to 1.0)	Depression‡	1139	38.4 (31.4 to 45.4)	Pain‡	1053	35.5 (27.8 to 43.1)	Fibromyalgia‡	604	20.3 (14.5 to 27.1)
Desvenlafaxine	897	29	3.2 (0.7 to8.0)	Depression‡	868	96.8 (91.9 to 99.3)	Anxiety disorder§	16	1.8 (0.0 to 5.5)	Menopausal hot flashes	6	0.7 (0.0 to 1.8)
Tricyclic antidepressants												
Amitriptyline	8993	8361	93.0 (89.6 to 95.7)	Pain	4349	48.4 (39.7 to 57.8)	Insomnia	2023	22.5 (13.6 to 31.3)	Migraine	1501	16.7 (12.2 to 21.9)
Doxepin	782	92	11.8 (3.2 to 21.3)	Insomnia‡	285	36.4 (24.1 to 49.3)	Depression‡	171	21.9 (9.7 to 35.9)	Anxiety disorder‡§	150	19.2 (9.0 to 32.0)
Nortriptyline	592	458	77.4 (59.9 to 89.5)	Pain	340	57.4 (35.0 to 74.1)	Depression‡	126	21.3 (9.5 to 38.5)	Anxiety disorder§	49	8.3 (2.4 to 20.0)
Trimipramine	562	165	29.4 (15.9 to 43.9)	Depression‡	397	70.6 (55.9 to 84.0)	Pain	93	16.5(5.2 to 29.9)	Insomnia	22	3.9 (0.2 to 11.3)
Imipramine	285	218	76.5 (55.4 to 90.6)	Panic disorder	69	24.2 (1.6 to 42.5)	Depression‡	67	23.5 (9.4 to 44.5)	Urinary disorders	64	22.5 (4.7 to 54.6)
Desipramine	216	127	58.8 (30.5 to 80.7)	Depression‡	89	41.2 (19.2 to 69.1)	Pain	81	37.5 (12.0 to 61.4)	Anxiety disorder§	16	7.4 (1.1 to 21.3)
Clomipramine	215	59	27.4 (8.9 to 51.1)	Depression‡	107	49.8 (25.8 to 72.3)	OCD‡	49	22.8 (7.2 to 42.0)	Anxiety disorder§	36	16.7 (2.6 to 37.9)
Other												
Trazodone	10070	8938	88.8 (81.5 to 93.7)	Insomnia	8303	82.5 (74.5 to 88.1)	Depression‡	1132	11.2 (6.3 to 18.4)	Anxiety disorder§	574	5.7 (3.8 to 8.2)
Bupropion	8384	780	9.3 (6.3 to 12.7)	Depression‡	7052	84.1 (79.9 to 87.9)	Nicotine dependence‡	565	6.7 (4.4 to 9.6)	ADHD	372	4.4 (2.5 to 6.9)
Mirtazapine	5908	622	10.5 (6.7 to 14.9)	Depression‡	5286	89.5 (85.1 to 93.7)	Anxiety disorder§	473	8.0 (4.3 to 13.2)	Insomnia	157	2.7 (0.8 to 4.8)

OCD=obsessive compulsive disorder; ADHD=attention deficit/hyperactivity disorder.
*Calculated using total number of prescriptions for the drug as the denominator.
†Calculated by a cluster bootstrap approach[26] to account for non-independence of prescriptions from the same physician and for the same patient. Reported 95% confidence intervals correspond to values at the 2.5th and 97.5th percentiles of the distribution of respective estimates across 1000 bootstrap re-samples.[26]
‡Indications approved for drug by Health Canada or the US Food and Drug Administration as of September 2015 (end of study period).
§Includes anxiety, generalised anxiety disorder, and other anxiety disorders. Excludes panic disorder, phobias, OCD, and post-traumatic stress disorder.

were supported by strong evidence. Radley and colleagues[29] combined antidepressants with anxiolytic and antipsychotic drugs, but again reported a similar off-label prescribing rate of 31%. However, the proportion of off-label prescriptions with strong scientific support in this study was notably lower than ours at only 6%, possibly due to the inclusion of other psychiatric drugs or because evidence to support some off-label antidepressant uses had not been generated at the time of the analysis. Finally, Walton and colleagues[25] presented results for only five antidepressants but similarly found that amitriptyline and trazodone were the antidepressants most frequently prescribed for off-label indications. However, their off-label prescribing rate was notably lower for amitriptyline (69%) and trazodone (43%) than our rates, possibly reflecting inter-country differences in the use of antidepressants versus other drugs to treat pain and insomnia.

In all of these studies, none of the authors assessed the proportion of off-label antidepressant prescriptions where the prescribed drug did not have strong evidence but another antidepressant from the same class existed that had strong evidence for the respective indication.

Potential explanations for off-label prescribing

Several contextual factors could contribute to physicians prescribing antidepressants for off-label indications. Firstly, the vast and increasing number of drugs on the market makes it challenging for physicians to keep track of which indications are approved for specific products,[30] especially when pharmaceutical companies have been known to promote drug use for off-label indications.[31] Secondly, constraints such as the list of drugs included on patients' health plan formularies could influence which drugs physicians prescribe, especially if physicians presume that drugs in the same class are interchangeable.[32 33] For example, in our setting, escitalopram was not covered for patients enrolled in the public drug insurance plan. We found that when study physicians prescribed SSRIs to patients with public drug insurance, they infrequently prescribed escitalopram (4.7% of all SSRI prescriptions for patients with public drug insurance) but frequently prescribed citalopram (51.4%). However, for patients with private drug insurance, study physicians equally prescribed escitalopram and citalopram (29.3% and 31.7% of all SSRI prescriptions for patients with private drug insurance, respectively).

Thirdly, primary care physicians might prescribe antidepressants off-label because alternative treatments for a given indication are contraindicated or perceived as higher risk medications. For example, benzodiazepines and Z drugs such as zolpidem and zaleplon have been shown to be efficacious for treating insomnia.[34] However, these drugs have been labelled as potentially inappropriate treatments for older adults, and if prescribed, could even negatively affect providers' quality and performance measures.[35] Many physicians who are concerned about the health of their older patients might consequently prescribe trazodone instead because they believe it is a safer treatment.

Finally, many off-label indications for antidepressants are symptom based conditions for which few approved drug treatments exist. Primary care physicians could be struggling to find effective treatments for these conditions and thus prescribe antidepressants as a last resort, indicating a gap in needed pharmacotherapy.

Implications of findings

For both primary care physicians and specialists (since specialists could initiate antidepressant treatment that is then continued by a primary care physician), our findings emphasise the importance of considering the level of evidence supporting risk-benefit when prescribing an antidepressant, especially if the drug is known to have important adverse side effects.[36] When evidence to support efficacy is lacking, physicians should exercise caution, prescribe conservatively, and inform patients of this information via a shared decision making process.[36] This ideal, however, is challenging to achieve because physicians face time constraints, the drug market and scientific literature are vast and ever-evolving, and many physicians find it challenging to critically appraise and interpret the results of epidemiological studies.[37] Indication based e-prescribing systems that are integrated with clinical decision support tools could help overcome these obstacles by notifying physicians when drugs are being prescribed off-label without supporting evidence and providing them with access to concise, up-to-date summaries of the available evidence. Providing the public with access to patient friendly resources about the level of scientific evidence supporting different treatment options for a given indication could further facilitate the decision making process between physicians and patients.

Our finding that among off-label prescriptions, 40% were for indications where the prescribed drug did not have strong evidence but another drug in the same class was approved or supported by strong evidence is clinically important. Many physicians might view this type of off-label prescribing as different from off-label prescribing without scientific evidence for the entire class because they assume that drugs within the same class are interchangeable.[38 39] However, class effects cannot be assumed because even slight differences in chemical structure between drugs can alter their pharmacodynamic and pharmacokinetic properties, leading to clinically relevant differences in efficacy and risk.[39] For example, statins have been shown to differ not only in efficacy[40] but also in safety, as demonstrated by the withdrawal of cerivastatin from the market in 2001 because the risk of rhabdomyolysis was 10 times higher for cerivastatin than other statins.[41] Clinical guidelines recommend that when physicians select a particular drug to prescribe, they should consider the level of scientific evidence supporting the specific drug.[42] It should not be assumed that all drugs within a class are likely to be efficacious for treating an indication when one member of the class has proven efficacy.

Finally, more evidence is needed on the clinical outcomes associated with antidepressant use for off-label indications. However, within a context of limited

doi: 10.1136/bmj.j603 | BMJ 2017;356:j603 | the bmj

resources, it is unlikely that randomised clinical trials will be conducted for each off-label drug-indication pair, especially for older drugs that are no longer owned by an innovator company.[9] Thus, in addition to randomised clinical trials, post-market drug surveillance systems represent valuable resources for assessing off-label antidepressant use. Such systems face challenges associated with measuring treatment indications and patient reported outcomes, but these challenges could be overcome by increasing the use of indication based e-prescribing systems and electronic health records that track patient outcomes. Indeed, this study demonstrates the benefits that indication based prescribing can have towards addressing knowledge gaps around off-label antidepressant prescribing.

Strengths and limitations

The key strength of this study is that it included more than 12 years of antidepressant prescriptions from an e-prescribing system where physicians systematically documented treatment indications at the point of prescribing. However, study participants were from one Canadian province where prescribers were generally younger and patients were generally older with more health complexities.[43] These characteristics could influence the generalisability of our findings, because younger physicians are more likely to prescribe drugs off-label without scientific evidence, and patients with more health complexities are less likely to receive off-label prescriptions.[7]

Another study strength is that physicians were unlikely to have altered their true responses when recording indications in the e-prescribing system because the dropdown menu did not distinguish between on-label and off-label indications for a drug. On the other hand, we could not identify when physicians consciously prescribed antidepressants off-label. Indeed, a portion of antidepressants in this study might have been prescribed off-label for a specific reason (eg, patient experienced side effects to another drug in the same class, or formulary restrictions).

Study considerations

Firstly, our estimates of off-label antidepressant prescribing were conservative because we did not consider other aspects of off-label drug use (eg, dose, frequency, duration of treatment), and we used the approved indications and available evidence at the end of the study period. Secondly, we presumed that approved indications for drugs were backed by strong scientific evidence, which might not have been true in some cases given that the quality of clinical trial evidence used by regulatory agencies as the basis for approving new therapeutics and supplemental indications has been shown to vary widely.[44 45]

Thirdly, to identify evidence based off-label uses for antidepressants, we used pre-established criteria that has been used in other studies.[7 18 25] However, our list of evidence based antidepressants for each indication might not always be identical to the recommendations from clinical guidelines. For example, recommendations from two national guidelines for managing anxiety related disorders[42 46] are similar but slightly more inclusive than ours. Finally, because regulatory bodies in North America and Europe are not entirely harmonised in their list of approved indications for drugs, slight discrepancies in the rate of off-label antidepressant use could exist between North America and Europe.

Conclusions

By using information from an indication based e-prescribing system, we found that when primary care physicians prescribed antidepressants for off-label indications, the prescribed drug was usually not supported by strong evidence for the respective indication. However, there was often another drug in the same class with strong evidence that could have been considered. These findings highlight an urgent need to produce more evidence on the risks and benefits of off-label antidepressant use and to provide physicians with this evidence at the point of prescribing. Technologies such as indication based e-prescribing systems and electronic health records have the potential to become essential components of effective post-market drug surveillance systems for monitoring and evaluating off-label antidepressant use. By integrating these technologies with knowledge databases and clinical decision support tools, they could also provide an effective means for communicating evidence back to physicians to optimise prescribing decisions.

We thank Claude Dagenais (academic adviser, Faculty of Pharmacy, University of Montréal) for reviewing the manuscript and providing substantive comments.

Contributors: JW extracted and had full access to all of the study data and takes responsibility for the integrity of the data and the accuracy of the data analysis. JW contributed to the study design; analysis and interpretation of the data; drafting of the manuscript; and critical revision of the manuscript for important intellectual content. AM and RT contributed to the study design; analysis and interpretation of the data; and critical revision of the manuscript for important intellectual content. MA and TE contributed to the analysis and interpretation of the data; and critical revision of the manuscript for important intellectual content. DB contributed to the interpretation of the data and critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript. JW is the guarantor.

Funding: This study was funded by the Canadian Institutes of Health Research (CIHR; grant IOP-112675). JW is supported by the Vanier Canada Graduate Scholarship (CIHR) and the Max E Binz Fellowship (Faculty of Medicine, McGill University). MA is a James McGill professor of biostatistics at McGill University. The funders had no role in the study design; collection, analysis, interpretation of the data; writing of the manuscript; or in the decision to submit the manuscript for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: support from the Canadian Institutes of Health Research for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: This study was approved by the McGill institutional review board.

Data sharing: No additional data available.

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported, and that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

1 Moore M, Yuen HM, Dunn N, Mullee MA, Maskell J, Kendrick T. Explaining the rise in antidepressant prescribing: a descriptive study using the general practice research database. *BMJ* 2009;339:b3999. doi:10.1136/bmj.b3999.
2 Prescribing and Medicines Team, Health and Social Care Information Centre. Prescriptions dispensed in the community: England 2005-2015. 2016. https://digital.nhs.uk/catalogue/PUB20664/pres-disp-com-eng-2005-15-rep.pdf
3 Hemels MEH, Koren G, Einarson TR. Increased use of antidepressants in Canada: 1981-2000. *Ann Pharmacother* 2002;36:1375-9. doi:10.1345/aph.1A331.
4 National Center for Health Statistics. Health, United States, 2010 with special feature on death and dying. 2011. https://www.cdc.gov/nchs/data/hus/hus10.pdf
5 Stone KJ, Viera AJ, Parman CL. Off-label applications for SSRIs. *Am Fam Physician* 2003;68:498-504.
6 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;2011:768-76.
7 Eguale T, Buckeridge DL, Winslade NE, Benedetti A, Hanley JA, Tamblyn R. Drug, patient, and physician characteristics associated with off-label prescribing in primary care. *Arch Intern Med* 2012;172:781-8. doi:10.1001/archinternmed.2012.340.
8 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;315:2230-2. doi:10.1001/jama.2016.3445.
9 Dresser R, Frader J. Off-label prescribing: a call for heightened professional and government oversight. *J Law Med Ethics* 2009;37:476-86, 396. doi:10.1111/j.1748-720X.2009.00408.x.
10 O'Malley PG. What does off-label prescribing really mean?*Arch Intern Med* 2012;172:759-60.
11 Hu XH, Bull SA, Hunkeler EM, et al. Incidence and duration of side effects and those rated as bothersome with selective serotonin reuptake inhibitor treatment for depression: patient report versus physician estimate. *J Clin Psychiatry* 2004;65:959-65. doi:10.4088/JCP.v65n0712.
12 Cascade E, Kalali AH, Kennedy SH. Real-world data on SSRI antidepressant side effects. *Psychiatry (Edgmont)* 2009;6:16-8.
13 Carvalho AF, Sharma MS, Brunoni AR, Vieta E, Fava GA. The safety, tolerability and risks associated with the use of newer generation antidepressant drugs: a critical review of the literature. *Psychother Psychosom* 2016;85:270-88. doi:10.1159/000447034.
14 Eom C-S, Lee H-K, Ye S, Park SM, Cho K-H. Use of selective serotonin reuptake inhibitors and risk of fracture: a systematic review and meta-analysis. *J Bone Miner Res* 2012;27:1186-95. doi:10.1002/jbmr.1554.
15 Anglin R, Yuan Y, Moayyedi P, Tse F, Armstrong D, Leontiadis GI. Risk of upper gastrointestinal bleeding with selective serotonin reuptake inhibitors with or without concurrent nonsteroidal anti-inflammatory use: a systematic review and meta-analysis. *Am J Gastroenterol* 2014;109:811-9. doi:10.1038/ajg.2014.82.
16 Dall M, Schaffalitzky de Muckadell OB, Lassen AT, Hansen JM, Hallas J. An association between selective serotonin reuptake inhibitor use and serious upper gastrointestinal bleeding. *Clin Gastroenterol Hepatol* 2009;7:1314-21. doi:10.1016/j.cgh.2009.08.019.
17 Wittich CM, Burkle CM, Lanier WL. Ten common questions (and their answers) about off-label drug use. *Mayo Clin Proc* 2012;87:982-90. doi:10.1016/j.mayocp.2012.04.017.
18 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;176:55-63. doi:10.1001/jamainternmed.2015.6058.
19 Pomey M-P, Forest P-G, Palley HA, Martin E. Public/private partnerships for prescription drug coverage: policy formulation and outcomes in Quebec's universal drug insurance program, with comparisons to the Medicare prescription drug program in the United States. *Milbank Q* 2007;85:469-98. doi:10.1111/j.1468-0009.2007.00496.x.
20 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;13:148-59. doi:10.1197/jamia.M1887.
21 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;33:559-67. doi:10.2165/11534580-000000000-00000.
22 Vigilance Santé. www.vigilance.ca/en/
23 Thomson Micromedex. Drugdex system (internet database). Greenwood Village, Colo. https://micromedex.com/compendia
24 Center for Medicare Advocacy. CMA report: Medicare coverage for off-label drug use. 2010. www.medicareadvocacy.org/cma-report-medicare-coverage-for-off-label-drug-use/
25 Walton SM, Schumock GT, Lee K-V, Alexander GC, Meltzer D, Stafford RS. Prioritizing future research on off-label prescribing: results of a quantitative evaluation. *Pharmacotherapy* 2008;28:1443-52. doi:10.1592/phco.28.12.1443.
26 Xiao Y, Abrahamowicz M. Bootstrap-based methods for estimating standard errors in Cox's regression analyses of clustered event times. *Stat Med* 2010;29:915-23. doi:10.1002/sim.3807.
27 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;67:972-82. doi:10.4088/JCP.v67n0615.
28 Rost K, Smith R, Matthews DB, Guise B. The deliberate misdiagnosis of major depression in primary care. *Arch Fam Med* 1994;3:333-7. doi:10.1001/archfami.3.4.333.
29 Radley DC, Finkelstein SN, Stafford RS. Off-label prescribing among office-based physicians. *Arch Intern Med* 2006;166:1021-6. doi:10.1001/archinte.166.9.1021.
30 Chen DT, Wynia MK, Moloney RM, Alexander GCUS. U.S. physician knowledge of the FDA-approved indications and evidence base for commonly prescribed drugs: results of a national survey. *Pharmacoepidemiol Drug Saf* 2009;18:1094-100. doi:10.1002/pds.1825.
31 Ghinea N, Lipworth W, Kerridge I. Off-label promotion of prescription medicine is it ever justifiable?*Ther Innov Regul Sci* 2015;49:359-63doi:10.1177/2168479015570337.
32 Sbarbaro JA. Can we influence prescribing patterns?*Clin Infect Dis* 2001;33(Suppl 3):S240-4. doi:10.1086/321856.
33 Stafford RS. Regulating off-label drug use--rethinking the role of the FDA. *N Engl J Med* 2008;358:1427-9. doi:10.1056/NEJMp0802107.
34 Wilson SJ, Nutt DJ, Alford C, et al. British Association for Psychopharmacology consensus statement on evidence-based treatment of insomnia, parasomnias and circadian rhythm disorders. *J Psychopharmacol* 2010;24:1577-601. doi:10.1177/0269881110379307.
35 Campanelli CM. American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc* 2012;60:616-31. doi:10.1111/j.1532-5415.2012.03923.x.
36 Largent EA, Miller FG, Pearson SD. Going off-label without venturing off-course: evidence and ethical off-label prescribing. *Arch Intern Med* 2009;169:1745-7. doi:10.1001/archinternmed.2009.314.
37 Godwin M, Seguin R. Critical appraisal skills of family physicians in Ontario, Canada. *BMC Med Educ* 2003;3:10. doi:10.1186/1472-6920-3-10.
38 McAlister FA, Laupacis A, Wells GA, Sackett DL. Evidence-Based Medicine Working Group. Users' Guides to the Medical Literature: XIX. Applying clinical trial results B. Guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA* 1999;282:1371-7. doi:10.1001/jama.282.14.1371.
39 Johnston A, Stafylas P, Stergiou GS. Effectiveness, safety and cost of drug substitution in hypertension. *Br J Clin Pharmacol* 2010;70:320-34. doi:10.1111/j.1365-2125.2010.03681.x.
40 Dieleman JP, van Wyk JT, van Wijk MAM, et al. Differences between statins on clinical endpoints: a population-based cohort study. *Curr Med Res Opin* 2005;21:1461-8. doi:10.1185/030079905X61866.
41 Furberg CD, Pitt B. Withdrawal of cerivastatin from the world market. *Curr Control Trials Cardiovasc Med* 2001;2:205-7. doi:10.1186/CVM-2-5-205.
42 Baldwin DS, Anderson IM, Nutt DJ, et al. Evidence-based pharmacological treatment of anxiety disorders, post-traumatic stress disorder and obsessive-compulsive disorder: a revision of the 2005 guidelines from the British Association for Psychopharmacology. *J Psychopharmacol* 2014;28:403-39. doi:10.1177/0269881114525674.
43 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 Bioeth* 2005;11:145-59doi:10.4067/S1726-569X2005000200005.
44 Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA* 2014;311:368-77. doi:10.1001/jama.2013.282034.
45 Wang B, Kesselheim AS. Characteristics of efficacy evidence supporting approval of supplemental indications for prescription drugs in United States, 2005-14: systematic review. *BMJ* 2015;351:h4679. doi:10.1136/bmj.h4679.
46 Katzman MA, Bleau P, Blier P, et al. Canadian Anxiety Guidelines Initiative Group on behalf of the Anxiety Disorders Association of Canada/Association Canadienne des troubles anxieux and McGill University. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry* 2014;14(Suppl 1):S1. doi:10.1186/1471-244X-14-S1-S1.

References

1. Government of Canada SC. Prescription medication use by Canadians aged 6 to 79 [Internet]. 2014 [cited 2015 May 5]. Available from: http://www.statcan.gc.ca/pub/82-003-x/2014006/article/14032-eng.htm

2. Canadian Institute for Health Information. Health Care Cost Drivers: The Facts. 2011.

3. (CIHI) CI for HI. Prescribed Drug Spending in Canada 2012 Report (formerly Drug Expenditure in Canada) [Internet]. 2014 [cited 2015 May 5]. Available from: https://secure.cihi.ca/estore/productFamily.htm?locale=en&pf=PFC2499

4. World Health Organization. International Drug Monitoring: The Role of National Centres. Report of a WHO Meeting. [Internet]. Geneva, Switzerland; 1972. Report No.: World Health Organization Technical Report Series No. 498. Available from: http://apps.who.int/iris/bitstream/10665/40968/1/WHO_TRS_498.pdf

5. Canadian Institute for Health Information. Adverse Drug Reaction-Related Hospitalizations Among Seniors, 2006 to 2011 [Internet]. 2013 Mar. Available from: https://secure.cihi.ca/free_products/Hospitalizations%20for%20ADR-ENweb.pdf

6. Wu C, Bell CM, Wodchis WP. Incidence and Economic Burden of Adverse Drug Reactions among Elderly Patients in Ontario Emergency Departments. Drug Saf. 2012;35(9):769–81.

7. Sultana J, Cutroneo P, Trifirò G. Clinical and economic burden of adverse drug reactions. J Pharmacol Pharmacother. 2013 Dec;4(Suppl1):S73–7.

8. Arthur N. The Importance of Pharmacovigilance: Safety Monitoring of medicinal products. United Kingdom: The World Health Organization; 2002.

9. Eguale T, Buckeridge DL, Winslade NE, Benedetti A, Hanley JA, Tamblyn R. Drug, patient, and physician characteristics associated with off-label prescribing in primary care. Arch Intern Med. 2012 May 28;172(10):781–8.

10. Radley DC, Finkelstein SN, Stafford RS. Off-label prescribing among office-based physicians. Arch Intern Med. 2006 May 8;166(9):1021–6.

11. Schiff GD, Seoane-Vazquez E, Wright A. Incorporating Indications into Medication Ordering--Time to Enter the Age of Reason. N Engl J Med. 2016 Jul 28;375(4):306–9.

12. Hemels MEH, Koren G, Einarson TR. Increased use of antidepressants in Canada: 1981-2000. Ann Pharmacother. 2002 Sep;36(9):1375–9.

13. OECD. Pharmaceutical consumption. In: Health at a Glance [Internet]. Organisation for Economic Co-operation and Development; 2013 [cited 2015 May 5]. p. 102–3. Available from: http://www.oecd-ilibrary.org/content/chapter/health_glance-2013-41-en

14. Mojtabai R, Olfson M. Proportion of antidepressants prescribed without a psychiatric diagnosis is growing. Health Aff Proj Hope. 2011 Aug;30(8):1434–42.

15. Stone KJ, Viera AJ, Parman CL. Off-label applications for SSRIs. Am Fam Physician. 2003 Aug 1;68(3):498–504.

16. 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 Oct;38(5):1311–7.

17. Gardarsdottir H, Heerdink ER, van Dijk L, Egberts ACG. Indications for antidepressant drug prescribing in general practice in the Netherlands. J Affect Disord. 2007 Feb;98(1–2):109–15.

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

19. Asnis GM, Thomas M, Henderson MA. Pharmacotherapy Treatment Options for Insomnia: A Primer for Clinicians. Int J Mol Sci [Internet]. 2015 Dec 30 [cited 2016 Jul 8];17(1). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4730295/

20. Mendelson W. A review of the evidence for the efficacy and safety of trazodone in insomnia. J Clin Psychiatry. 2005 Apr;66(4):469–76.

21. Khouzam HR. A review of trazodone use in psychiatric and medical conditions. Postgrad Med. 2017 Jan;129(1):140–8.

22. Bossini L, Casolaro I, Koukouna D, Cecchini F, Fagiolini A. Off-label uses of trazodone: a review. Expert Opin Pharmacother. 2012 Aug;13(12):1707–17.

23. Management of Chronic Insomnia Disorder in Adults: A Clinical Practice Guideline From the American College of Physicians. Ann Intern Med. 2016 Jul 19;165(2):I-26.

24. Carvalho AF, Sharma MS, Brunoni AR, Vieta E, Fava GA. The Safety, Tolerability and Risks Associated with the Use of Newer Generation Antidepressant Drugs: A Critical Review of the Literature. Psychother Psychosom. 2016;85(5):270–88.

25. Cascade E, Kalali AH, Kennedy SH. Real-World Data on SSRI Antidepressant Side Effects. Psychiatry Edgmont. 2009 Feb;6(2):16–8.

26. Coupland C, Dhiman P, Morriss R, Arthur A, Barton G, Hippisley-Cox J. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. BMJ. 2011;343:d4551.

27. Wilkes S. The use of bupropion SR in cigarette smoking cessation. Int J Chron Obstruct Pulmon Dis. 2008 Mar;3(1):45–53.

28. Doyle JT, Dawber TR, Kannel WB, Heslin AS, Kahn HA. Cigarette Smoking and Coronary Heart Disease. N Engl J Med. 1962 Apr 19;266(16):796–801.

29. 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;2011:768–76.

30. Larson MJ, Miller K, Fleming KJ. Treatment with antidepressant medications in private health plans. Adm Policy Ment Health. 2007 Mar;34(2):116–26.

31. 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 Jun;26(6):1343–53.

32. 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 Sep;21(9):980–8.

33. Simon GE, Stewart C, Beck A, Ahmedani BK, Coleman KJ, Whitebird RR, et al. National prevalence of receipt of antidepressant prescriptions by persons without a psychiatric diagnosis. Psychiatr Serv Wash DC. 2014 Jul;65(7):944–6.

34. 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 Jun;67(6):972–82.

35. Davis KAS, Sudlow CLM, Hotopf M. Can mental health diagnoses in administrative data be used for research? A systematic review of the accuracy of routinely collected diagnoses. BMC Psychiatry [Internet]. 2016 Jul 26 [cited 2017 Jul 7];16. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4960739/

36. Tamblyn R, Huang A, Kawasumi Y, Bartlett G, Grad R, Jacques A, et al. The Development and Evaluation of an Integrated Electronic Prescribing and Drug Management System for Primary Care. J Am Med Inform Assoc JAMIA. 2006;13(2):148–59.

37. 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 Jul 1;33(7):559–67.

38. Kruppa J, Liu Y, Biau G, Kohler M, König IR, Malley JD, et al. Probability estimation with machine learning methods for dichotomous and multicategory outcome: theory. Biom J Biom Z. 2014 Jul;56(4):534–63.

39. van der Laan MJ, Polley EC, Hubbard AE. Super learner. Stat Appl Genet Mol Biol. 2007;6:Article25.

40. Polley E, LeDell E, Kennedy C, Lendle S, Laan M van der. SuperLearner: Super Learner Prediction. R package version 2.0-21. [Internet]. 2016. Available from: https://CRAN.R-project.org/package=SuperLearner

41. Hubbard A, Munoz ID, Decker A, Holcomb JB, Schreiber MA, Bulger EM, et al. Time-Dependent Prediction and Evaluation of Variable Importance Using SuperLearning in High Dimensional Clinical Data. J Trauma Acute Care Surg. 2013 Jul;75(1 0 1):S53–60.

42. Petersen ML, LeDell E, Schwab J, Sarovar V, Gross R, Reynolds N, et al. Super learner analysis of electronic adherence data improves viral prediction and may provide strategies for selective HIV RNA monitoring. J Acquir Immune Defic Syndr 1999. 2015 May 1;69(1):109–18.

43. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der LAAN MJ. Mortality prediction in the ICU: can we do better? Results from the Super ICU Learner Algorithm (SICULA) project, a population-based study. Lancet Respir Med. 2015 Jan;3(1):42–52.

44. Rose S. Mortality risk score prediction in an elderly population using machine learning. Am J Epidemiol. 2013 Mar 1;177(5):443–52.

45. Neugebauer R, Fireman B, Roy JA, Raebel MA, Nichols GA, O'Connor PJ. Super learning to hedge against incorrect inference from arbitrary parametric assumptions in marginal structural modeling. J Clin Epidemiol. 2013 Aug;66(8 0):S99–109.

46. Luo G. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. Netw Model Anal Health Inform Bioinforma. 2016 Dec 1;5(1):18.

47. Kreula E, Hemminki E. Frequency of psychotropic drug prescribing for children in Tampere, Finland. Acta Paediatr Scand. 1978 Jul;67(4):449–52.

48. Mercier A, Auger-Aubin I, Lebeau J-P, Schuers M, Boulet P, Van Royen P, et al. Why do general practitioners prescribe antidepressants to their patients? A pilot study. Biopsychosoc Med. 2014;8:17.

49. Bergman U, Dahlström M, Gunnarsson C, Westerholm B. Why are psychotropic drugs prescribed to out-patients? A methodological study. Eur J Clin Pharmacol. 1979 May 21;15(4):249–56.

50. Henriksson S, Boëthius G, Håkansson J, Isacsson G. Indications for and outcome of antidepressant medication in a general population: a prescription database and medical record study, in Jämtland county, Sweden, 1995. Acta Psychiatr Scand. 2003 Dec;108(6):427–31.

51. Chermá MD, Löfgren U-B, Almkvist G, Hallert C, Bengtsson F. Assessment of the prescription of antidepressant drugs in elderly nursing home patients: a clinical and laboratory follow-up investigation. J Clin Psychopharmacol. 2008 Aug;28(4):424–31.

52. Mackay FJ, Dunn NR, Wilton LV, Pearce GL, Freemantle SN, Mann RD. A comparison of fluvoxamine, fluoxetine, sertraline and paroxetine examined by observational cohort studies. Pharmacoepidemiol Drug Saf. 1997 Jul;6(4):235–46.

53. Harris T, Carey IM, Shah SM, DeWilde S, Cook DG. Antidepressant Prescribing in Older Primary Care Patients in Community and Care Home Settings in England and Wales. J Am Med Dir Assoc. 2012 Jan;13(1):41–7.

54. Hassan L, Senior J, Webb RT, Frisher M, Tully MP, While D, et al. Prevalence and appropriateness of psychotropic medication prescribing in a nationally representative cross-sectional survey of male and female prisoners in England. BMC Psychiatry. 2016 Oct 10;16(1):346.

55. Meijer WE, Heerdink ER, Pepplinkhuizen LP, van Eijk JT, Leufkens HG. Prescribing patterns in patients using new antidepressants. Br J Clin Pharmacol. 2001 Feb;51(2):181–3.

56. Volkers AC, Heerdink ER, van Dijk L. Antidepressant use and off-label prescribing in children and adolescents in Dutch general practice (2001-2005). Pharmacoepidemiol Drug Saf. 2007 Sep;16(9):1054–62.

57. Ruths S, Straand J, Nygaard HA. Psychotropic drug use in nursing homes--diagnostic indications and variations between institutions. Eur J Clin Pharmacol. 2001 Sep;57(6–7):523–8.

58. Buhl Sørensen C, Bøhm Jepsen E, Thomsen PH, Dalsgaard S. Indications for and use of antidepressants in child and adolescent psychiatry--a cross-sectional survey in Denmark. Eur Child Adolesc Psychiatry. 2003 Jun;12(3):114–21.

59. 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 Jan;18(1):7–15.

60. Steinhausen H-C, Bisgaard C. Nationwide time trends in dispensed prescriptions of psychotropic medication for children and adolescents in Denmark. Acta Psychiatr Scand. 2014 Mar;129(3):221–31.

61. Noordam R, Aarts N, Verhamme KM, Sturkenboom MCM, Stricker BH, Visser LE. Prescription and indication trends of antidepressant drugs in the Netherlands between 1996 and 2012: a dynamic population-based study. Eur J Clin Pharmacol. 2015 Mar;71(3):369–75.

62. Petty DR, House A, Knapp P, Raynor T, Zermansky A. Prevalence, duration and indications for prescribing of antidepressants in primary care. Age Ageing. 2006 Sep 1;35(5):523–6.

63. Haw C, Stubbs J. Off-label psychotropic prescribing for young persons in medium security. J Psychopharmacol Oxf Engl. 2010 Oct;24(10):1491–8.

64. John A, Marchant AL, Fone DL, McGregor JI, Dennis MS, Tan JOA, et al. Recent trends in primary-care antidepressant prescribing to children and young people: an e-cohort study. Psychol Med. 2016 Dec;46(16):3315–27.

65. Sarginson J, Webb RT, Stocks SJ, Esmail A, Garg S, Ashcroft DM. Temporal trends in antidepressant prescribing to children in UK primary care, 2000–2015. J Affect Disord. 2017 Mar 1;210:312–8.

66. Trifirò G, Barbui C, Spina E, Moretti S, Tari M, Alacqua M, et al. Antidepressant drugs: prevalence, incidence and indication of use in general practice of Southern Italy during the years 2003-2004. Pharmacoepidemiol Drug Saf. 2007 May;16(5):552–9.

67. Zullino DF, Schwartz B, Bilancioni R, Baumann P. Off-label utilization of antidepressants. Acta Medica (Hradec Kralove). 2008;51(1):19–24.

68. Bourgeois J, Elseviers MM, Van Bortel L, Petrovic M, Vander Stichele RH. The use of antidepressants in Belgian nursing homes: focus on indications and dosages in the PHEBE study. Drugs Aging. 2012 Sep;29(9):759–69.

69. Abbing-Karahagopian V, Huerta C, Souverein PC, Abajo F de, Leufkens HGM, Slattery J, et al. Antidepressant prescribing in five European countries: application of common definitions to assess the prevalence, clinical observations, and methodological implications. Eur J Clin Pharmacol. 2014 Jul 1;70(7):849–57.

70. Loosbrock DL, Tomlin ME, Robinson RL, Obenchain RL, Croghan TW. Appropriateness of prescribing practices for serotonergic antidepressants. Psychiatr Serv Wash DC. 2002 Feb;53(2):179–84.

71. Pomerantz JM, Finkelstein SN, Berndt ER, Poret AW, Walker LE, Alber RC, et al. Prescriber intent, off-label usage, and early discontinuation of antidepressants: a retrospective physician survey and data analysis. J Clin Psychiatry. 2004 Mar;65(3):395–404.

72. Larson M, Miller K, Fleming KJ. Antidepressant Medication Use in Private Insurance Health Plans, 2002 [Internet]. PubMed Journals. [cited 2017 Apr 19]. Available from: https://ncbi.nlm.nih.gov/labs/articles/16452692/

73. Cascade EF, Kalali AH, Thase ME. Use of Antidepressants. Psychiatry Edgmont. 2007 Dec;4(12):25–8.

74. Mark TL. For What Diagnoses Are Psychotropic Medications Being Prescribed? CNS Drugs. 2010 Apr 1;24(4):319–26.

75. Lee E, Teschemaker AR, Johann-Liang R, Bazemore G, Yoon M, Shim K-S, et al. Off-label prescribing patterns of antidepressants in children and adolescents. Pharmacoepidemiol Drug Saf. 2012 Feb;21(2):137–44.

76. O'Brien PL, Cummings N, Mark TL. Off-Label Prescribing of Psychotropic Medication, 2005–2013: An Examination of Potential Influences. Psychiatr Serv. 2017 Jan 17;appi.ps.201500482.

77. Patten SB, Esposito E, Carter B. Reasons for antidepressant prescriptions in Canada. Pharmacoepidemiol Drug Saf. 2007 Jul 1;16(7):746–52.

78. Bridges-Webb C, Mant A, Hall W. Psychotropic Drug Prescribing in an Australian General Practice. Fam Pract. 1984 Jun 1;1(2):106–12.

79. Levin CA, Wei W, Akincigil A, Lucas JA, Bilder S, Crystal S. Prevalence and treatment of diagnosed depression among elderly nursing home residents in Ohio. J Am Med Dir Assoc. 2007 Nov;8(9):585–94.

80. 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):145–59.

81. Breton M, Lévesque J-F, Pineault R, Hogg W. Primary Care Reform: Can Quebec's Family Medicine Group Model Benefit from the Experience of Ontario's Family Health Teams? Healthc Policy. 2011 Nov;7(2):e122–35.

82. 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 Jan;15(1):19–29.

83. Vigilance Santé [Internet]. [cited 2015 Jun 13]. Available from: http://www.vigilance.ca/en/

84. Thomson Micromedex. Drugdex system (internet database). Greenwood Village, Colo;

85. 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 Nov;51(11):938–42.

86. National Center for Health Statistics. Health, United States, 2010 With Special Feature on Death and Dying [Internet]. Hyattsville, MD; 2011 [cited 2015 Jul 23]. Available from: http://www.cdc.gov/nchs/data/hus/hus10.pdf

87. Carey V, Zeger SL, Diggle P. Modelling Multivariate Binary Data with Alternating Logistic Regressions. Biometrika. 1993 Sep 1;80(3):517–26.

88. Keller DL. In Defense of Off-label Prescribing. JAMA Intern Med. 2016 Jun 1;176(6):861.

89. Moore M, Yuen HM, Dunn N, Mullee MA, Maskell J, Kendrick T. Explaining the rise in antidepressant prescribing: a descriptive study using the general practice research database. BMJ. 2009 Oct 15;339:b3999.

90. Prescribing and Medicines Team, Health and Social Care Information Centre. Prescriptions Dispensed in the Community: England 2005-2015 [Internet]. 2016 Jul [cited 2016 Sep 14]. Available from: http://digital.nhs.uk/catalogue/PUB20664/pres-disp-com-eng-2005-15-rep.pdf

91. 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 May 24;315(20):2230–2.

92. Dresser R, Frader J. Off-Label Prescribing: A Call for Heightened Professional and Government Oversight. J Law Med Ethics J Am Soc Law Med Ethics. 2009;37(3):476–396.

93. O'Malley PG. What does off-label prescribing really mean? Arch Intern Med. 2012 May 28;172(10):759–60.

94. Hu XH, Bull SA, Hunkeler EM, Ming E, Lee JY, Fireman B, et al. Incidence and duration of side effects and those rated as bothersome with selective serotonin reuptake inhibitor treatment for depression: patient report versus physician estimate. J Clin Psychiatry. 2004 Jul;65(7):959–65.

95. Eom C-S, Lee H-K, Ye S, Park SM, Cho K-H. Use of selective serotonin reuptake inhibitors and risk of fracture: a systematic review and meta-analysis. J Bone Miner Res Off J Am Soc Bone Miner Res. 2012 May;27(5):1186–95.

96. Anglin R, Yuan Y, Moayyedi P, Tse F, Armstrong D, Leontiadis GI. Risk of upper gastrointestinal bleeding with selective serotonin reuptake inhibitors with or without concurrent nonsteroidal anti-inflammatory use: a systematic review and meta-analysis. Am J Gastroenterol. 2014 Jun;109(6):811–9.

97. Dall M, Schaffalitzky de Muckadell OB, Lassen AT, Hansen JM, Hallas J. An association between selective serotonin reuptake inhibitor use and serious upper gastrointestinal bleeding. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc. 2009 Dec;7(12):1314–21.

98. Wittich CM, Burkle CM, Lanier WL. Ten Common Questions (and Their Answers) About Off-label Drug Use. Mayo Clin Proc. 2012 Oct;87(10):982–90.

99. 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. 2015 Nov 2;1–9.

100. Pomey M-P, Forest P-G, Palley HA, Martin E. Public/Private Partnerships for Prescription Drug Coverage: Policy Formulation and Outcomes in Quebec's Universal Drug Insurance Program, with Comparisons to the Medicare Prescription Drug Program in the United States. Milbank Q. 2007 Sep;85(3):469–98.

101. Center for Medicare Advocacy. CMA REPORT: MEDICARE COVERAGE FOR OFF-LABEL DRUG USE [Internet]. [cited 2016 Dec 14]. Available from: http://www.medicareadvocacy.org/cma-report-medicare-coverage-for-off-label-drug-use/

102. Walton SM, Schumock GT, Lee K-V, Alexander GC, Meltzer D, Stafford RS. Prioritizing future research on off-label prescribing: results of a quantitative evaluation. Pharmacotherapy. 2008 Dec;28(12):1443–52.

103. Xiao Y, Abrahamowicz M. Bootstrap-based methods for estimating standard errors in Cox's regression analyses of clustered event times. Stat Med. 2010 Mar 30;29(7–8):915–23.

104. Rost K, Smith R, Matthews DB, Guise B. The deliberate misdiagnosis of major depression in primary care. Arch Fam Med. 1994 Apr;3(4):333–7.

105. Chen DT, Wynia MK, Moloney RM, Alexander GC. U.S. physician knowledge of the FDA-approved indications and evidence base for commonly prescribed drugs: results of a national survey. Pharmacoepidemiol Drug Saf. 2009 Nov;18(11):1094–100.

106. Ghinea N, Lipworth W, Kerridge I. Off-Label Promotion of Prescription Medicine Is It Ever Justifiable? Ther Innov Regul Sci. 2015 May 1;49(3):359–63.

107. Sbarbaro JA. Can We Influence Prescribing Patterns? Clin Infect Dis. 2001 Sep 15;33(Supplement 3):S240–4.

108. Stafford RS. Regulating Off-Label Drug Use — Rethinking the Role of the FDA. N Engl J Med. 2008 Apr 3;358(14):1427–9.

109. Wilson SJ, Nutt DJ, Alford C, Argyropoulos SV, Baldwin DS, Bateson AN, et al. British Association for Psychopharmacology consensus statement on evidence-based treatment of insomnia, parasomnias and circadian rhythm disorders. J Psychopharmacol Oxf Engl. 2010 Nov;24(11):1577–601.

110. Campanelli CM. American Geriatrics Society Updated Beers Criteria for Potentially Inappropriate Medication Use in Older Adults. J Am Geriatr Soc. 2012 Apr;60(4):616–31.

111. Largent EA, Miller FG, Pearson SD. Going Off-label Without Venturing Off-Course: Evidence and Ethical Off-label Prescribing. Arch Intern Med. 2009 Oct 26;169(19):1745–7.

112. Godwin M, Seguin R. Critical appraisal skills of family physicians in Ontario, Canada. BMC Med Educ. 2003;3:10.

113. McAlister FA, Laupacis A, Wells GA, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature: Xix. applying clinical trial resultsb. guidelines for determining whether a drug is exerting (more than) a class effect. JAMA. 1999 Oct 13;282(14):1371–7.

114. Johnston A, Stafylas P, Stergiou GS. Effectiveness, safety and cost of drug substitution in hypertension. Br J Clin Pharmacol. 2010 Sep;70(3):320–34.

115. Dieleman JP, Wyk JT van, Wijk MAM van, Herpen G van, Straus SMJM, Dunselman H, et al. Differences between statins on clinical endpoints: a population-based cohort study. Curr Med Res Opin. 2005 Sep 1;21(9):1461–8.

116. Furberg CD, Pitt B. Withdrawal of cerivastatin from the world market. Curr Control Trials Cardiovasc Med. 2001;2(5):205–7.

117. Baldwin DS, Anderson IM, Nutt DJ, Allgulander C, Bandelow B, den Boer JA, et al. Evidence-based pharmacological treatment of anxiety disorders, post-traumatic stress disorder and obsessive-compulsive disorder: a revision of the 2005 guidelines from the British Association for Psychopharmacology. J Psychopharmacol Oxf Engl. 2014 May;28(5):403–39.

118. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. JAMA. 2014 Jan 22;311(4):368–77.

119. Wang B, Kesselheim AS. Characteristics of efficacy evidence supporting approval of supplemental indications for prescription drugs in United States, 2005-14: systematic review. BMJ. 2015 Sep 23;351:h4679.

120. Katzman MA, Bleau P, Blier P, Chokka P, Kjernisted K, Van Ameringen M. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. BMC Psychiatry. 2014 Jul 2;14(Suppl 1):S1.

121. 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 Feb 21;356:j603.

122. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects--advantages and disadvantages. Nat Clin Pract Rheumatol. 2007 Dec;3(12):725–32.

123. Cadieux G, Buckeridge DL, Jacques A, Libman M, Dendukuri N, Tamblyn R. Patient, physician, encounter, and billing characteristics predict the accuracy of syndromic surveillance case definitions. BMC Public Health. 2012;12:166.

124. Bossini L, Coluccia A, Casolaro I, Benbow J, Amodeo G, De Giorgi R, et al. Off-Label Trazodone Prescription: Evidence, Benefits and Risks. Curr Pharm Des. 2015;21(23):3343–51.

125. Carroll DG, Kelley KW. Use of antidepressants for management of hot flashes. Pharmacotherapy. 2009 Nov;29(11):1357–74.

126. Slatkoff J, Greenfield B. Pharmacological treatment of attention-deficit/hyperactivity disorder in adults. Expert Opin Investig Drugs. 2006 Jun;15(6):649–67.

127. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005 Nov;43(11):1130–9.

128. Lacasse A, Ware MA, Dorais M, Lanctôt H, Choinière M. Is the Quebec provincial administrative database a valid source for research on chronic non-cancer pain? Pharmacoepidemiol Drug Saf. 2015 Sep;24(9):980–90.

129. Wu C-H, Tung Y-C, Lin T-K, Chai C-Y, Su Y-F, Tsai T-H, et al. Hip Fracture in People with Erectile Dysfunction: A Nationwide Population-Based Cohort Study. PloS One. 2016;11(4):e0153467.

130. Birtwhistle R, Morkem R, Peat G, Williamson T, Green ME, Khan S, et al. Prevalence and management of osteoarthritis in primary care: an epidemiologic cohort study from the Canadian Primary Care Sentinel Surveillance Network. CMAJ Open. 2015 Sep;3(3):E270-275.

131. Widdifield J, Bombardier C, Bernatsky S, Paterson JM, Green D, Young J, et al. An administrative data validation study of the accuracy of algorithms for identifying rheumatoid arthritis: the influence of the reference standard on algorithm performance. BMC Musculoskelet Disord. 2014;15:216.

132. Neogi T. The Epidemiology and Impact of Pain in Osteoarthritis. Osteoarthr Cartil OARS Osteoarthr Res Soc. 2013 Sep;21(9):1145–53.

133. Lee YC. Effect and Treatment of Chronic Pain in Inflammatory Arthritis. Curr Rheumatol Rep. 2013 Jan;15(1):300.

134. (CIHI) CI for HI. Conversion Tables (for use with ICD-10-CA/CCI) [Internet]. [cited 2016 Jul 27]. Available from: https://secure.cihi.ca/estore/productSeries.htm?pc=PCC85

135. Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol. 2008;56(1):45–50.

136. Leisenring W, Pepe MS. Regression Modelling of Diagnostic Likelihood Ratios for the Evaluation of Medical Diagnostic Tests. Biometrics. 1998 Jun 1;54(2):444–52.

137. van Walraven C, English S, Austin PC. Administrative database code accuracy did not vary notably with changes in disease prevalence. J Clin Epidemiol [Internet]. [cited 2016 Aug 16]; Available from: http://www.sciencedirect.com/science/article/pii/S0895435616301470

138. McGee S. Simplifying Likelihood Ratios. J Gen Intern Med. 2002 Aug;17(8):647–50.

139. Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring, III JR. Medical Epidemiology, Fourth Edition. The McGraw-Hill Companies, Inc.; 2005.

140. Topitz A, Benda N, Saumer G, Friedrich F, König D, Soulier N, et al. [Prevalence and recognition of depression among inpatients of non-psychiatric hospital departments].

Neuropsychiatr Klin Diagn Ther Rehabil Organ Ges Österr Nervenärzte Psychiater. 2015;29(2):63–70.

141. Wittchen H-U, Mühlig S, Beesdo K. Mental disorders in primary care. Dialogues Clin Neurosci. 2003 Jun;5(2):115–28.

142. Arnold LM, Clauw DJ, Dunegan LJ, Turk DC. A Framework for Fibromyalgia Management for Primary Care Providers. Mayo Clin Proc. 2012 May;87(5):488–96.

143. Shepherd G, Mohorn P, Yacoub K, May DW. Adverse drug reaction deaths reported in United States vital statistics, 1999-2006. Ann Pharmacother. 2012 Feb;46(2):169–75.

144. Lazarou J, Pomeranz BH, Corey PN. Incidence of Adverse Drug Reactions in Hospitalized Patients: A Meta-analysis of Prospective Studies. JAMA. 1998 Apr 15;279(15):1200–5.

145. Bonn D. Adverse drug reactions remain a major cause of death. The Lancet. 1998 Apr 18;351(9110):1183.

146. Kumar A. Pharmacovigilance: Importance, concepts, and processes. Am J Health Syst Pharm. 2017 Feb 24;ajhp151031.

147. Pratt L, Brody D, Gu Q. Antidepressant use in persons aged 12 and over: United States, 2005-2008. NCHS data brief, no 76. [Internet]. National Center for Health Statistics; 2011 [cited 2015 May 5]. Available from: http://www.cdc.gov/nchs/data/databriefs/db76.htm

148. Wong J, Abrahamowicz M, Buckeridge DL, Tamblyn R. Assessing the accuracy of using diagnostic codes from administrative data to determine antidepressant treatment indications: a validation study (unpublished).

149. Medical services | RAMQ [Internet]. [cited 2017 Mar 11]. Available from: http://www.ramq.gouv.qc.ca/en/citizens/health-insurance/healthcare/Pages/medical-services.aspx

150. McCall WV. Off-label Use of Prescription Medications for Insomnia: Sedating Antidepressants, Antipsychotics, Anxiolytics, and Anticonvulsants (Chapter 34). In: Insomnia: Diagnosis and Treatment. UK: Informa Healthcare; 2010. p. 397–409.

151. J. Katon W. Epidemiology and treatment of depression in patients with chronic medical illness. Dialogues Clin Neurosci. 2011 Mar;13(1):7–23.

152. Druss BG, Rask K, Katon WJ. Major Depression, Depression Treatment, and Quality of Primary Medical Care. Gen Hosp Psychiatry. 2008;30(1):20–5.

153. 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.

154. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. New York: Springer-Verlag; 2009.

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

156. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiol Camb Mass. 2010 Jan;21(1):128–38.

157. 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 Aug;50(12):3464–85.

158. 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 Jan 30;27(2):157-172; discussion 207-212.

159. Sutton DA, Moldofsky H, Badley EM. Insomnia and health problems in Canadians. Sleep. 2001 Sep 15;24(6):665–70.

160. Smith BH, Elliott AM, Chambers WA, Smith WC, Hannaford PC, Penny K. The impact of chronic pain in the community. Fam Pract. 2001 Jun;18(3):292–9.

161. Wong ST, Manca D, Barber D, Morkem R, Khan S, Kotecha J, et al. The diagnosis of depression and its treatment in Canadian primary care practices: an epidemiological study. CMAJ Open. 2014 Oct 1;2(4):E337–42.

162. Kop WJ, Seliger SL, Fink JC, Katz R, Odden MC, Fried LF, et al. Longitudinal Association of Depressive Symptoms with Rapid Kidney Function Decline and Adverse Clinical Renal Disease Outcomes. Clin J Am Soc Nephrol. 2011 Mar 10;CJN.03840510.

163. Nutt D, Wilson S, Paterson L. Sleep disorders as core symptoms of depression. Dialogues Clin Neurosci. 2008 Sep;10(3):329–36.

164. Berger A, Dukes E, Martin S, Edelsberg J, Oster G. Characteristics and healthcare costs of patients with fibromyalgia syndrome. Int J Clin Pract. 2007 Sep;61(9):1498–508.

165. Wilkins T, Jarvis K, Patel J. Diagnosis and management of Crohn's disease. Am Fam Physician. 2011 Dec 15;84(12):1365–75.

166. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer-Verlag; 2009.

167. Kruse CS, Goswamy R, Raval Y, Marawi S. Challenges and Opportunities of Big Data in Health Care: A Systematic Review. JMIR Med Inform [Internet]. 2016 Nov 21 [cited 2017 Jul 13];4(4). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5138448/

168. Westreich D, Lessler J, Funk MJ. Propensity score estimation: machine learning and classification methods as alternatives to logistic regression. J Clin Epidemiol. 2010 Aug;63(8):826–33.

169. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLOS ONE. 2017 Apr 4;12(4):e0174944.

170. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J. 2015;13:8–17.

171. Díaz I, Hubbard A, Decker A, Cohen M. Variable Importance and Prediction Methods for Longitudinal Problems with Missing Variables. PLoS ONE [Internet]. 2015 Mar 27 [cited 2017 Aug 5];10(3). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4376910/

172. Wong J, Abrahamowicz M, Buckeridge DL, Tamblyn R. Derivation and validation of a multivariable model to predict when primary care physicians prescribe antidepressants for indications besides depression. Unpublished. 2017;

173. Dasgupta A, Sun YV, König IR, Bailey-Wilson JE, Malley JD. Brief Review of Regression-Based and Machine Learning Methods in Genetic Epidemiology: The Genetic Analysis Workshop 17 Experience. Genet Epidemiol. 2011;35(Suppl 1):S5-11.

174. Foster KR, Koprowski R, Skufca JD. Machine learning, medical diagnosis, and biomedical engineering research - commentary. Biomed Eng OnLine. 2014 Jul 5;13:94.

175. Therneau T, Atkinson B, Ripley B. rpart: Recursive Partitioning and Regresson Trees. R package version 4.1-11. [Internet]. 2017. Available from: https://CRAN.R-project.org/parkage=rpart

176. Rose S. A Machine Learning Framework for Plan Payment Risk Adjustment. Health Serv Res. 2016 Dec;51(6):2358–74.

177. Hinton GE. How neural networks learn from experience. Sci Am. 1992 Sep;267(3):144–51.

178. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol. 1996 Nov;49(11):1225–31.

179. Yu W, Liu T, Valdez R, Gwinn M, Khoury MJ. Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. BMC Med Inform Decis Mak. 2010 Mar 22;10:16.

180. Zanaty EA. Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification. Egypt Inform J. 2012 Nov 1;13(3):177–83.

181. Christmann A, Steinwart I. Support Vector Machines. New York, NY: Springer Science+Business Media, LLC; 2008. (Information Science and Statistics).

182. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, [Internet]. Vienna, Austria; Available from: https://www.R-project.org/

183. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 1(33):1–22.

184. Liaw A, Wiener M. Classification and Regression by randomForest. R News. 2(3):18–22.

185. Venables W, Ripley B. Modern Applied Statistics with S. Fourth Edition. Springer, New York; 2002.

186. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.6-8. 2017.

187. Ballings M, Van den Poel D. AUC: Threshold independent performance measures for probabilistic classifiers. R package version 0.3.0. 2013; Available from: https://CRAN.R-project.org/package=AUC

188. Karim ME, Platt RW, BeAMS study group. Estimating inverse probability weights using super learner when weight-model specification is unknown in a marginal structural Cox model context. Stat Med. 2017 Jun 15;36(13):2032–47.

189. Schuler MS, Rose S. Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. Am J Epidemiol. 2017 Jan 1;185(1):65–73.

190. Li J, Handorf E, Bekelman J, Mitra N. Propensity score and doubly robust methods for estimating the effect of treatment on censored cost. Stat Med. 2016 May 30;35(12):1985–99.

191. Khondoker M, Dobson R, Skirrow C, Simmons A, Stahl D. A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies. Stat Methods Med Res. 2016 Oct 1;25(5):1804–23.

192. Ferguson JM. SSRI Antidepressant Medications: Adverse Effects and Tolerability. Prim Care Companion J Clin Psychiatry. 2001 Feb;3(1):22–7.

193. Sifferlin A. Half of the People Taking Antidepressants Aren't Depressed: Study [Internet]. Time. [cited 2017 Aug 12]. Available from: http://time.com/4345517/antidepressants-depression-insomnia-depression-migraine/

194. Almendrala A. 8 Surprising Reasons People Are Taking Antidepressants [Internet]. HuffPost Canada. 2016 [cited 2017 Aug 12]. Available from: http://www.huffingtonpost.com/entry/reasons-people-are-taking-antidepressants_us_57472146e4b03ede441411a6

195. Welsh A. Antidepressants used for more than depression: study [Internet]. 2016 [cited 2017 Aug 12]. Available from: http://www.cbsnews.com/news/nearly-half-of-antidepressants-not-prescribed-for-depression-study/

196. Mozes A. “Off-label” antidepressant use is common, but is it safe? [Internet]. [cited 2017 Aug 12]. Available from: http://www.cbsnews.com/news/are-off-label-antidepressant-prescriptions-safe/

197. International RC. Antidepressants prescribed for many people not depressed [Internet]. RCI | English. 2016 [cited 2017 Aug 12]. Available from: http://www.rcinet.ca/en/2016/05/26/antidepressants-off-label-study-mcgill/

198. Crowe K. Questioning “off-label” antidepressant use, and can you trust online doctor reviews: CBC’s health newsletter [Internet]. CBC News. [cited 2017 Aug 12]. Available from: http://www.cbc.ca/news/health/off-label-antidepressant-use-lacks-evidence-and-a-link-between-math-skills-and-your-health-1.3997685

199. Takahashi M. CTV News interview [Internet]. 2016. Available from: http://montreal.ctvnews.ca/video?playlistId=1.2920314

200. Derfel A. Nearly half of all antidepressants not prescribed for depression: Quebec study [Internet]. Montreal Gazette. 2016 [cited 2017 Aug 12]. Available from: http://montrealgazette.com/news/local-news/nearly-half-of-all-antidepressants-not-prescribed-for-depression-quebec-study

201. Dick F, Tevaearai H. Significance and Limitations of the p Value. Eur Soc Vasular Surg. 2015;50:815.

202. McMahon AD, MacDonald TM. Design issues for drug epidemiology. Br J Clin Pharmacol. 2000 Nov;50(5):419–25.

203. Pearson BR. Under-reporting of Adverse Drug Reactions: The Need for an Automated Reporting System. Rev Interdiscip Sci Santé - Interdiscip J Health Sci. 2016 Mar 10;3(1):15–20.

204. Hohl CM, Karpov A, Reddekopp L, Doyle-Waters M, Stausberg J. ICD-10 codes used to identify adverse drug events in administrative data: a systematic review. J Am Med Inform Assoc JAMIA. 2014 Jun;21(3):547–57.

205. Iqbal E, Mallah R, Jackson RG, Ball M, Ibrahim ZM, Broadbent M, et al. Identification of Adverse Drug Events from Free Text Electronic Patient Records and Information in a Large Mental Health Case Register. PLOS ONE. 2015 Aug 14;10(8):e0134208.

206. Gysbers M, Reichley R, Kilbridge PM, Noirot L, Nagarajan R, Dunagan WC, et al. Natural language processing to identify adverse drug events. AMIA Annu Symp Proc AMIA Symp. 2008 Nov 6;961.

207. Patel R, Lloyd T, Jackson R, Ball M, Shetty H, Broadbent M, et al. Mood instability and clinical outcomes in mental health disorders: A natural language processing (NLP) study. Eur Psychiatry. 2016 Mar 1;33:S224.

208. Carrell DS, Halgrim S, Tran D-T, Buist DSM, Chubak J, Chapman WW, et al. Using Natural Language Processing to Improve Efficiency of Manual Chart Abstraction in Research: The Case of Breast Cancer Recurrence. Am J Epidemiol. 2014 Mar 15;179(6):749–58.