CanJPsychiatry 2015;60(8):362–368

Original Research

Validation of a Population-Based Algorithm to Detect Chronic Psychotic Illness

Paul Kurdyak, MD, PhD[1]; Elizabeth Lin, PhD[2]; Diane Green, BSc[3]; Simone Vigod, MD, MSc[4]

[1] Director, Health Systems Research, Social and Epidemiological Research, Centre for Addiction and Mental Health, Toronto, Ontario; Lead, Mental Health and Addictions Research Program, Institute for Clinical Evaluative Sciences, Toronto, Ontario; Assistant Professor, Department of Psychiatry and Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario.
Correspondence: Centre for Addiction and Mental Health, 33 Russell Street, Toronto, ON M5S 2S1; paul.kurdyak@camh.ca.

[2] Research Scientist, Provincial System Support Program, Centre for Addiction and Mental Health, Toronto, Ontario; Adjunct Scientist, Institute for Clinical Evaluative Sciences, Toronto, Ontario; Associate Professor, Department of Psychiatry, University of Toronto, Toronto, Ontario.

[3] Analyst, Institute for Clinical Evaluative Sciences, Toronto, Ontario.

[4] Scientist, Women's College Research Institute, Women's College Hospital, Toronto, Ontario; Assistant Professor, Department of Psychiatry and Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario; Adjunct Scientist, Institute for Clinical Evaluative Sciences, Toronto, Ontario.

Key Words: schizophrenia, chart validation, sensitivity, specificity, administrative data, positive predictive value, negative predictive value

Received September 2014, revised, and accepted January 2015.

Objective: To validate algorithms to detect people with chronic psychotic illness in population-based health administrative databases.

Method: We developed 8 algorithms to detect chronic psychotic illness using hospitalization and physician service claims data from administrative health databases in Ontario to identify cases of chronic psychotic illness between 2002 and 2007. Diagnostic data abstracted from the records of 281 randomly selected psychiatric patients from 2 hospitals in Toronto were linked to the administrative data cohort to test sensitivity, specificity, and positive predictive values (PPV) and negative predictive values.

Results: Using only hospitalization data to capture chronic psychotic illness yielded the highest specificity (range 69.9% to 84.7%) and the highest PPV (range 55.2% to 80.8%). Using physician service claims in addition to hospitalization data to capture cases increased sensitivity (range 90.1% to 98.8%) but decreased specificity (range 31.1% to 68.0%) and PPV (range 38.4% to 71.1%).

Conclusion: Using health administrative data to study population-based outcomes for people with chronic psychotic illness is feasible and valid. Researchers can select case identification methods based on whether a more sensitive or more specific definition of chronic psychotic illness is desired.

Validation d'un algorithme dans la population pour détecter la maladie psychotique chronique

Objectif : Valider des algorithmes pour détecter les personnes souffrant de maladie psychotique chronique dans des bases de données populationnelles administratives de santé.

Méthode : Nous avons mis au point 8 algorithmes pour détecter la maladie psychotique chronique à l'aide de données d'hospitalisation et de facturation des médecins tirées des bases de données administratives de santé de l'Ontario, afin d'identifier les cas de maladie psychotique chronique entre 2002 et 2007. Les données diagnostiques obtenues des dossiers de 281 patients psychiatriques choisis au hasard dans 2 hôpitaux de Toronto ont été liées à la cohorte des données administratives pour vérifier la sensibilité, la spécificité, les valeurs prédictives positives (VPP) et les valeurs prédictives négatives.

Résultats : L'utilisation des données d'hospitalisation seulement pour trouver la maladie psychotique chronique donnait la spécificité la plus élevée (étendue de 69,9 % à 84,7 %) et les VPP les plus élevées (écart de 55,2 % à 80,8 %). L'utilisation des données de facturation des médecins en plus des données d'hospitalisation pour détecter les cas augmentait la sensibilité (éttendue de 90,1 % à 98,8 %) mais diminuait la spécificité (étendue de 31,1 % à 68,0 %) et les VPP (étendue de 38,4 % à 71,1 %).

Conclusion : L'utilisation de données administratives de santé pour étudier des résultats dans la population concernant les personnes souffrant de maladie psychotique chronique est faisable et valide. Les chercheurs peuvent choisir les méthodes d'identification de cas selon qu'ils préfèrent une définition plus sensible ou plus spécifique de la maladie psychotique chronique.

Schizophrenia and other chronic psychotic illnesses, such as schizoaffective disorder, are severe and disabling mental illnesses.[1] While these illnesses are relatively low in prevalence (0.5% to 1%),[2] they are associated with poor psychiatric[3] and medical outcomes, including mortality,[4–6] and impart a significant economic burden to the health care system.[7] A large literature indicates that poor access to timely and effective health care resources partly explains the poor outcomes in this population.[5,8–10] Studies of health care service use among people with chronic psychotic illness are necessary to understand why these poor outcomes occur, and to assess the effectiveness of system-level interventions that have been implemented to optimize health trajectories. Health administrative data are advantageous for these types of studies because they allow for population-level surveillance of risk factors and outcomes, as well as for the evaluation of a multitude of health services and supports for large numbers of people, at minimal cost.

Despite the frequent use of health administrative data for health services and outcome studies in this area, no definitive method for identifying people with chronic psychotic illnesses on a population-level has been recommended. This is in contrast to successful validation studies for cardiovascular illnesses[11,12] and diabetes[13] that have set the stage for important population-based outcomes research for these conditions. There have been previous efforts to use health administrative data to identify people with chronic psychotic illnesses within a population using US Medicaid data and Scandinavian health registries.[14,15] Using hospitalization data, as has been done previously,[15] accurately identifies people with schizophrenia, but will miss people who are treated primarily in ambulatory settings. This can be problematic for 2 reasons: the estimated prevalence of chronic psychotic illnesses will be lower than the true prevalence, and the sample of people will have a more severe form of illness than a population-based sample that includes people detected and treated in ambulatory care settings. Further, to our knowledge, no previous studies have considered the diagnosis of psychotic disorder NOS. This is a diagnosis that is often applied to people with chronic psychotic illnesses, particularly in ambulatory settings where clinicians do not have sufficient time, compared with clinicians working in inpatient settings, to ascertain whether criteria for schizophrenia or schizoaffective disorder have been met. There is evidence that most people who are initially diagnosed with nonspecific psychotic diagnoses like psychotic disorder NOS are ultimately diagnosed with schizophrenia or schizoaffective disorder.[16] A recent Canadian study[17] exploring diagnostic stability the year following a first-episode psychosis suggested that schizophrenia is the most stable diagnosis and other diagnoses tended to shift toward schizophrenia spectrum diagnoses. Accordingly, it is important to determine the impact of including or excluding psychotic disorder NOS when using administrative data to study health services and outcomes for people with chronic psychotic illness. It is equally important to restrict the patient population to people with diagnoses most like schizophrenia, and to exclude patients with delusional disorder and substance-related psychotic illnesses, because people with schizophrenia and related diagnoses represent a meaningfully homogeneous patient population from the perspective of diagnosis, prognosis, and service use.

Clinical Implications

- Health administrative data can be used to generate population-based case registries of people with chronic psychotic illnesses.
- Using hospitalization diagnoses will increase PPV, but may miss people with less severe forms of chronic psychotic illness who are never hospitalized.
- Using physician visit data in addition to hospitalization data increases the sensitivity of chronic psychotic illness algorithms at the cost of specificity and PPV.

Limitations

- The sample used to validate diagnostic algorithms was relatively small, limiting the power to detect small differences in test characteristics.
- Health records were abstracted from hospitals and not primary care settings which limits the generalizability of the findings to people with chronic psychotic illnesses who neither come into contact with specialists (psychiatrists) and (or) have had a psychiatric hospitalization.

Abbreviations

CAMH	Centre for Addiction and Mental Health
CIHI-DAD	Canadian Institutes of Health Information—Discharge Abstract Database
DSM	Diagnostic and Statistical Manual of Mental Disorders
ICD	International Classification of Diseases
ICES	Institute for Clinical Evaluative Sciences
NOS	not otherwise specified
NPV	negative predictive value
OHIP	Ontario Health Information Plan
OMHRS	Ontario Mental Health Reporting System
OPCRIT	Operational Criteria Checklist
PPV	positive predictive value
RPDB	Registered Persons Database

Aims of Our Study

Our study aimed to validate algorithms to detect people with chronic psychotic illnesses in Ontario in health administrative databases. We compared algorithms containing administrative data from inpatient hospitalization and physician service claims with diagnostic information abstracted directly from clinical records (our reference standard) to determine the most optimal method for identifying people with schizophrenia using health administrative data. We used 2 reference standards for chronic psychotic illness: schizophrenia and (or) schizoaffective disorder (specific definition); and schizophrenia, schizoaffective disorder and (or) psychosis NOS (nonspecific definition). Our hypotheses were 2-fold: including physician visits in addition to hospitalization diagnoses would increase sensitivity but decrease the PPV of algorithms to capture cases of chronic psychotic illness, and including psychotic disorder NOS as a diagnostic category in addition to using schizophrenia and schizoaffective disorder alone would increase the sensitivity without affecting specificity of diagnostic health administrative algorithms.

Methods

Study Design and Setting

We conducted a chart abstraction study in Ontario. The administrative health data used to create case definitions for identifying people with schizophrenia or schizoaffective disorder were obtained from ICES. ICES is a nonprofit research organization that uses population-level databases to evaluate health care services and their effectiveness in Ontario. Patient-level records in these data are anonymously linked to each other through a unique identifier (ICES key number) using the RPDB for every Ontario resident with an assigned health care number. Ontario has a single payer, universal health care system that provides coverage for health care, including physician visits and hospitalizations, and captures information on all of these visits. To validate the administrative health case definitions, clinical charts were abstracted at 2 hospitals in Toronto. One (CAMH) is an academic tertiary care psychiatric hospital providing long-term care and inpatient care for people with schizophrenia, as well as many other types of psychiatric care to people from Toronto and across Ontario; the other (St Joseph's Health Centre) is a community-based urban hospital with a general psychiatry department that provides psychiatric crisis service, as well as inpatient and ambulatory care mainly to people living in the Toronto area. These 2 hospitals were chosen to be representative of 2 different types of health care centres where people with psychiatric conditions, including schizophrenia, are likely to receive inpatient and ambulatory care. From each hospital, we aimed to abstract 150 charts restricted to patients with psychiatric diagnoses between 2002 and 2007. Nineteen charts were excluded because they did not contain sufficient health information to permit abstraction, leaving a total of 281 charts for the study sample.

Administrative Data Case Definitions

Data sources for the administrative data algorithms were the population-level health administrative databases at ICES. Psychiatric hospitalizations were captured from 2 databases. Prior to 2005, all psychiatric hospitalizations were captured in the CIHI-DAD, which used the ICD-10 diagnostic system from 2002 onward. In 2005, the OMHRS was created and captures all hospitalizations occurring in mental health–designated hospital beds, recording discharge diagnoses using the DSM-IV diagnostic system. Physician service claims were captured using the OHIP database for physician billings for physician visits and consultations that lists patient diagnostic codes. Using diagnostic codes for schizophrenia and schizoaffective disorder (the specific diagnostic codes: CIHI-DAD ICD-10 F20 and F25; OMHRS DSM-IV schizophrenia and schizoaffective disorder; and OHIP 295), we identified people with hospitalization only; hospitalization and 1 physician visit; hospitalization and 2 physician visits within 24 months; and hospitalization and 3 physician visits within 36 months. We increased the number of physician visit time frames for diagnosis to determine whether increasing the requirements for ambulatory-based diagnosis corresponded to an increase in specificity. We repeated the above algorithms to add psychotic disorder NOS (the nonspecific diagnostic codes: ICD 10 F29; DSM-IV; and OHIP 298) to the first algorithm containing schizophrenia and schizoaffective disorder based on evidence that people with this diagnosis are often indistinguishable from people with schizophrenia and schizoaffective disorder.[16] Thus we had 4 specific and 4 nonspecific algorithms, for a total of 8 administrative case definitions for chronic psychotic illness.

Chart Abstraction and Validation

The charts were abstracted using the OPCRIT, a software package used to systematically abstract health records to generate psychiatric diagnoses, including from DSM-IV (for Ontario psychiatric hospitalization diagnoses) and ICD-9 (for Ontario physician billing data). OPCRIT has good interrater reliability when compared with trained clinicians abstracting charts[18–20] and is now incorporated into the clinical record at major hospitals in the United Kingdom.[21] A mental health nurse was trained on the OPCRIT and abstracted information from all 281 charts to generate a specific diagnosis of schizophrenia or schizoaffective disorder, and a nonspecific diagnosis of psychotic disorder that also included the category of psychosis NOS. For validation, the 281 patient records (148 from CAMH and 133 from St Joseph's Health Centre) were linked to the health administrative data housed at ICES using patient health card numbers. After linkage, all patient identifiers were removed from the data and destroyed. The diagnoses generated from the various health administrative data algorithms were then compared with the chart-abstracted diagnoses (both the specific and nonspecific definitions) to produce sensitivity, specificity, PPV, and NPV for each algorithm.

Table 1 Characteristics of patients at the 2 abstraction hospitals

Variable	CAMH *n* = 148	St Joseph's Health Centre *n* = 133	*P*
Age at index, years			
Mean (SD)	40.66 (15.32)	40.44 (15.11)	0.90
Median (IQR)	41 (20)	39 (21)	0.99
Sex on RPDB, *n* (%)			0.03
Female	66 (44.6)	77 (57.9)	
Male	82 (55.4)	56 (42.1)	
Income quintile, *n* (%)			0.23
1	36 (24.3)	50 (37.6)	
2	40 (27.0)	28 (21.1)	
3	23 (15.5)	15 (11.3)	
4	21 (14.2)	20 (15.0)	
5	24 (16.2)	18 (13.5)	
Specific diagnoses, *n* (%)			<0.001
0	119 (80.4)	77 (57.9)	
1	29 (19.6)	56 (42.1)	
Nonspecific diagnoses, *n* (%)			<0.001
0	97 (65.5)	53 (39.8)	
1	51 (34.5)	80 (60.2)	

CAMH = Centre for Addiction and Mental Health; RPDB = Registered Persons Database

Demographic Variables

For the purposes of describing the sample, we captured age, sex, and neighbourhood income quintile for all subjects. Age and sex are available from the RPDB. Neighbourhood income is derived using the patient's postal code information and neighbourhood income from Census data.

This project was approved by the Research Ethics Boards at CAMH, St Joseph's Health Centre, and Sunnybrook Health Sciences Centre.

Results

Study Cohort

Among the 281 patients, 85 (30%) were identified as having schizophrenia or schizoaffective disorder using OPCRIT. The number increased to 130 (46%) when the OPCRIT definition included a diagnosis of psychotic disorder NOS. There were more cases of chronic psychotic illness at St Joseph's Health Centre, the community-based hospital, than CAMH. Characteristics of the study population are described in Table 1.

Test Characteristics of the Health Administrative Data Algorithms

Test characteristics for each of the 8 health administrative algorithms, compared with both the specific and nonspecific OPCRIT diagnoses, are described in Table 2. For the specific OPCRIT diagnosis (that is, schizophrenia or schizoaffective disorder only), the sensitivity of the health administrative data algorithms ranged from 82.4% to 98.8%, and specificity ranged from 31.1% to 74.5%. The PPVs were moderate, ranging from 38.4% to 58.3%, whereas the NPV was much higher, ranging from 90.7% to 99.0%. For the nonspecific OPCRIT diagnosis (schizophrenia, schizoaffective disorder, or psychotic disorder NOS), the sensitivity of the health administrative algorithms was similar to the OPCRIT-specific diagnosis, ranging from 74.0% to 96.0%, but the specificity was higher, ranging from 38.0% to 84.7%. While the addition of physician service claims in all cases improved sensitivity, health administrative data algorithms using hospitalization data only generated the highest specificity and PPV. For health administrative algorithms that included physician visits, specificity was increased with an increasing number of visits, with the specificity reaching 90% and NPV reaching 89%, but with a modest decrease in specificity (68%) and PPV (71%), compared with algorithms using only hospitalization data. There were no meaningful differences between the health administrative algorithms that incorporated the psychosis NOS diagnosis, compared with those that did not.

Discussion

We used OPCRIT, a well-tested instrument developed to derive psychiatric diagnoses from chart abstraction, to test whether we could generate a valid, population-based sample of people with chronic psychotic illness within a

Table 2 Test characteristics

Characteristic	Sensitivity, % OR (95% CI)	Specificity, % OR (95% CI)	PPV, % OR (95% CI)	NPV, % OR (95% CI)
OPCRIT-specific				
Administrative codes for schizophrenia and schizoaffective disorder				
Hospitalization	82.4 (74.2 to 90.5)	74.5 (68.4 to 80.6)	58.3 (49.5 to 67.2)	90.7 (86.2 to 95.2)
Hospitalization or 1 MD visit	97.6 (94.4 to 100.0)	40.8 (33.9 to 47.7)	41.7 (34.9 to 48.6)	97.6 (94.2 to 100.0)
Hospitalization or 2 MD visits within 24 months	97.6 (94.4 to 100.0)	51.5 (44.5 to 58.5)	46.6 (39.3 to 54.0)	98.1 (95.4 to 100.0)
Hospitalization or 3 MD visits within 36 months	96.5 (92.5 to 100.0)	57.1 (50.2 to 64.1)	49.4 (41.8 to 57.0)	97.4 (94.5 to 100.0)
Administrative codes for schizophrenia, schizoaffective disorder, and psychosis NOS				
Hospitalization	87.1 (79.9 to 94.2)	69.9 (63.5 to 76.3)	55.6 (47.2 to 64.1)	92.6 (88.3 to 96.8)
Hospitalization or 1 MD visit	98.8 (96.5 to 100.0)	31.1 (24.6 to 37.6)	38.4 (31.9 to 44.8)	98.4 (95.3 to 100.0)
Hospitalization or 2 MD visits within 24 months	98.8 (96.5 to 100.0)	41.8 (34.9 to 48.7)	42.4 (35.5 to 49.3)	98.8 (96.4 to 100.0)
Hospitalization or 3 MD visits within 36 months	98.8 (96.5 to 100.0)	49.0 (42.0 to 56.0)	45.7 (38.5 to 52.8)	99.0 (97.0 to 100.0)
OPCRIT-nonspecific				
Administrative codes for schizophrenia and schizoaffective disorder				
Hospitalization	74.0 (66.5 to 81.6)	84.7 (78.9 to 90.4)	80.8 (73.8 to 87.9)	78.9 (72.6 to 85.2)
Hospitalization or 1 MD visit	94.7 (90.8 to 98.5)	50.0 (42.0 to 58.0)	62.3 (55.6 to 69.0)	91.5 (85.4 to 97.5)
Hospitalization or 2 MD visits within 24 months	91.6 (86.9 to 96.4)	61.3 (53.5 to 69.1)	67.4 (60.5 to 74.3)	89.3 (83.4 to 95.3)
Hospitalization or 3 MD visits within 36 months	90.1 (85.0 to 95.2)	68.0 (60.5 to 75.5)	71.1 (64.2 to 78.0)	88.7 (82.9 to 94.5)
Administrative codes for schizophrenia, schizoaffective disorder, and psychosis NOS				
Hospitalization	79.4 (72.5 to 86.3)	80.7 (74.3 to 87.0)	78.2 (71.2 to 85.2)	81.8 (75.5 to 88.0)
Hospitalization or 1 MD visit	96.2 (92.9 to 99.5)	38.0 (30.2 to 45.8)	57.5 (51.0 to 64.1)	91.9 (85.2 to 98.7)
Hospitalization or 2 MD visits within 24 months	93.9 (89.8 to 98.0)	50.0 (42.0 to 58.0)	62.1 (55.4 to 68.9)	90.4 (84.0 to 96.7)
Hospitalization or 3 MD visits within 36 months	93.1 (88.8 to 97.5)	58.7 (50.8 to 66.5)	66.3 (59.5 to 73.1)	90.7 (84.9 to 96.5)

MD = medical doctor; NOS = not otherwise specified; NPV = negative predictive value; OPCRIT = Operational Criteria Checklist; PPV = positive predictive value

large, population-based health administrative database. We tested 2 hypotheses: including physician visits in addition to hospitalization diagnoses would increase sensitivity but decrease the PPV of algorithms to capture cases of chronic psychotic illness; and including psychotic disorder NOS as a diagnostic category in addition to using schizophrenia and schizoaffective disorder alone would increase the sensitivity without affecting specificity of diagnostic health administrative algorithms. Our first hypothesis was borne out; using hospitalization data from health administrative records, we were able to adequately identify people with chronic psychotic illness (schizophrenia, schizoaffective disorder, and [or] psychosis NOS) with sensitivity of 74%, specificity of 85%, PPV of 80%, and NPV of 79%. Using a health administrative data definition of a hospitalization or 3 physician service claims within 3 years, we were able to increase the sensitivity of the diagnosis to 90% and the NPV to 89% with a modest decrease in specificity (68%) and PPV (71%). Our second hypothesis was also true; including psychotic disorder NOS to the 2 main primary psychotic illnesses (schizophrenia and schizoaffective disorder) increased sensitivity without decreasing specificity. The specificity described in our study is likely less than the true specificity when applied to the entire population. The prevalence of psychiatric illnesses is much lower in the general population, thus the rate of false

positives will likely be much lower, and the specificity will be higher than we reported. While investigators may choose to use algorithms that incorporate physician service claims, our study has demonstrated that health administrative data can, and should, be used to study population-based outcomes for people with chronic psychotic illnesses such as schizophrenia and schizoaffective disorder. It has also demonstrated that health administrative data codes for schizophrenia and schizoaffective disorder capture people with these disorders, as well as those with the more nonspecific psychosis NOS designation.

There have been many publications using health administrative data to describe population-based outcomes for people with psychosis. Surprisingly, there have been relatively few validation studies to support the use of health administrative data for such studies.[14,15] These studies are quite dated[14] or are restricted to hospitalization data only.[15] Restricting data to hospitalizations creates the risk of not accounting for less severe cases of schizophrenia and thus biasing the population-based registry to people who have worse outcomes. Our study reveals that including cases ascertained through physician visit data increases sensitivity, but at the risk of generating false positives (that is, it lowers PPV). Thus whether one includes cases through physician billings or restricts to those only acquired through hospitalization diagnosis will depend on whether a researcher wants a chronic psychosis cohort that is inclusive but has false-positive cases or a chronic psychosis cohort that is highly specific but misses cases, particularly people who may have less severe forms of chronic psychosis. Therefore, if a researcher wants to be more certain of the diagnosis, and is interested in a sample that tends to be more severely ill, then using hospitalizations only for case detection would be advantageous, with the disadvantage of a relatively high false-negative rate. Conversely, if a researcher wants to capture as close to the entire population of people with chronic psychotic illnesses as possible, then including physician visits for case detection would be advantageous, with the disadvantage of a relatively high false-positive rate. Both types of algorithm have merits and shortcomings, and one approach may be to use one algorithm (for example, hospitalization only) and then conduct a sensitivity analysis using the other in a given study. There has also been variation in the diagnoses used to capture people with chronic psychotic illnesses.[14,22–24] We tested 2 health administrative data algorithms: one that included only schizophrenia and schizoaffective disorder, and one that added psychotic disorder NOS to the 2 specific diagnoses. Including psychotic disorder NOS generally improved the test characteristics of the health administrative data algorithms relative to the OPCRIT reference standard, suggesting that, in Ontario, many psychiatrists use the psychotic disorder NOS diagnosis for people who would be defined as having schizophrenia or schizoaffective disorder by OPCRIT criteria.

The strengths of our research are as follows: we used a well-described and tested instrument (OPCRIT) to abstract health records; we used 2 different settings (a tertiary care psychiatric hospital and a community-based hospital) to increase the generalizability of our findings; we specifically tested different sets of diagnoses to capture people with chronic psychotic illnesses; and we included physician visit data to attempt to capture a more representative population-based sample of people with chronic psychotic illnesses. We also constructed diagnostic algorithms that are easily replicable in other jurisdictions and, as such, are likely generalizable to any clinical setting where health administrative data are routinely collected. There are also limitations to our analysis. First, our sample size was relatively small; however, we were able to generate test characteristics with reasonable confidence intervals. Second, we used hospitals as sites for abstracting health records. We elected to not conduct a chart abstraction study in primary care settings for 2 reasons: the prevalence of schizophrenia is quite low, making it very unlikely that we would be able to develop a validated sample of subjects with chronic psychosis from primary care setting chart abstraction; and with the high rate of hospitalization and the high volume of ambulatory care and early intervention programs run by hospitals, people with chronic psychotic illnesses are likely to have contact (and therefore generate either hospitalization or physician visit health administrative data) at community hospitals or tertiary care mental health centres.

In summary, we have demonstrated that diagnostic algorithms using health administrative data can be used to capture population-based cohorts of people with chronic psychotic illnesses. We have further demonstrated that using an inclusive set of diagnoses captures a wider range of severity as demonstrated by an improved sensitivity. Further, using physician visits as an option to detect cases also yields a wider range of severity with a much higher sensitivity, but with a decrease in specificity. The algorithm one uses depends on the need for representativeness (at the cost of including false-positive cases) or a more specific cohort (at the cost of missing potentially less severely ill cases). Validating algorithms to identify people with chronic psychotic illnesses is a very important and foundational step in the goal of conducting high quality research mapping the outcomes of this vulnerable population.

Acknowledgements

Our study was funded by a grant from the Ontario Mental Health Foundation. This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results, and conclusions reported in our paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred.

Dr Kurdyak designed the study, oversaw data collection and analysis, and drafted the manuscript. Dr Lin contributed to the study design and drafting of the protocol. Dr Gnam drafted the original protocol and contributed to the study design. Diane Green and Dr Vigod contributed to the study

design and the statistical analysis. All authors contributed to, and approved, the final manuscript.

The authors declare no conflicts of interest.

We thank Dr Ryan Todd for his assistance with chart abstraction.

References

1. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 4th ed. Text revision. Washington (DC): APA; 2000.
2. Goldner EM, Hsu L, Waraich P, et al. Prevalence and incidence studies of schizophrenic disorders: a systematic review of the literature. Can J Psychiatry. 2002;47(9):833–843.
3. Jobe TH, Harrow M. Long-term outcome of patients with schizophrenia: a review. Can J Psychiatry. 2005;50(14):892–900.
4. Bresee LC, Majumdar SR, Patten SB, et al. Prevalence of cardiovascular risk factors and disease in people with schizophrenia: a population-based study. Schizophr Res. 2010;117(1):75–82.
5. Brown S, Inskip H, Barraclough B. Causes of the excess mortality of schizophrenia. Br J Psychiatry. 2000;177:212–217.
6. Brown S, Kim M, Mitchell C, et al. Twenty-five year mortality of a community cohort with schizophrenia. Br J Psychiatry. 2010;196(2):116–121.
7. Ratnasingham S, Cairney J, Manson H, et al. The burden of mental illness and addiction in Ontario. Can J Psychiatry. 2013;58(9):529–537.
8. Druss BG, Bradford WD, Rosenheck RA, et al. Quality of medical care and excess mortality in older patients with mental disorders. Arch Gen Psychiatry. 2001;58(6):565–572.
9. Kurdyak P, Vigod S, Calzavara A, et al. High mortality and low access to care following incident acute myocardial infarction in individuals with schizophrenia. Schizophr Res. 2012;142(1–3):52–57.
10. Copeland LA, Zeber JE, Wang CP, et al. Patterns of primary care and mortality among patients with schizophrenia or diabetes: a cluster analysis approach to the retrospective study of healthcare utilization. BMC Health Serv Res. 2009;9:127.
11. Austin P, Daly P, Tu J. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. Am Heart J. 2002;144(2):290–296.
12. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. Med Care. 2005;43(2):182–188.
13. Hux JE, Ivis F, Flintoft V, et al. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. Diabetes Care. 2002;25(3):512–516.
14. Lurie N, Popkin M, Dysken M, et al. Accuracy of diagnoses of schizophrenia in Medicaid claims. Hosp Community Psychiatry. 1992;43(1):69–71.
15. Pihlajamaa J, Suvisaari J, Henriksson M, et al. The validity of schizophrenia diagnosis in the Finnish Hospital Discharge Register: findings from a 10-year birth cohort sample. Nord J Psychiatry. 2008;62(3):198–203.
16. Veen ND, Selten JP, Schols D, et al. Diagnostic stability in a Dutch psychosis incidence cohort. Br J Psychiatry. 2004;185:460–464. doi: 10.1192/bjp.185.6.460.
17. Pope MA, Joober R, Malla AK. Diagnostic stability of first-episode psychotic disorders and persistence of comorbid psychiatric disorders over 1 year. Can J Psychiatry. 2013;58(10):588–594.
18. Craddock M, Asherson P, Owen MJ, et al. Concurrent validity of the OPCRIT diagnostic system. Comparison of OPCRIT diagnoses with consensus best-estimate lifetime diagnoses. Br J Psychiatry. 1996;169(1):58–63.
19. Williams J, Farmer AE, Ackenheil M, et al. A multicentre inter-rater reliability study using the OPCRIT computerized diagnostic system. Psychol Med. 1996;26(4):775–783.
20. Brittain PJ, Stahl D, Rucker J, et al. A review of the reliability and validity of OPCRIT in relation to its use for the routine clinical assessment of mental health patients. Int J Methods Psychiatr Res. 2013;22(2):110–137. doi: 10.1002/mpr.1382.
21. Rucker J, Newman S, Gray J, et al. OPCRIT+: an electronic system for psychiatric diagnosis and data collection in clinical and research settings. Br J Psychiatry. 2011;199(2):151–155. doi: 10.1192/bjp.bp.110.082925.
22. Goldner EM, Jones W, Waraich P. Using administrative data to analyze the prevalence and distribution of schizophrenic disorders. Psychiatr Serv. 2003;54(7):1017–1021.
23. Ekholm B, Ekholm A, Adolfsson R, et al. Evaluation of diagnostic procedures in Swedish patients with schizophrenia and related psychoses. Nord J Psychiatry. 2005;59(6):457–464.
24. Vares M, Ekholm A, Sedvall GC, et al. Characterization of patients with schizophrenia and related psychoses: evaluation of different diagnostic procedures. Psychopathology. 2006;39(6):286–295. doi: 10.1159/000095733.