Clinical Interventions in Aging

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

# Validity of the Updated Rx-Risk Index as a Disease Identification and Risk-Adjustment Tool for Use in Observational Health Studies

Imaina Widagdo[1], Mhairi Kerr[1], Lisa Kalisch Ellett[1], Clement Schlegel[2], Elham Sadeqzadeh[2], Alvin Wang[2], Allison Louise Clarke[2], Nicole Pratt[1]

[1]Quality Use of Medicines and Pharmacy Research Centre, Clinical and Health Sciences, University of South Australia, Adelaide, South Australia, Australia; [2]Data and Analytics Branch, Health Economics and Research Division, Department of Health and Aged Care, Canberra, Australian Capital Territory, Australia

Correspondence: Imaina Widagdo, University of South Australia, GPO Box 2471, Adelaide, SA, 5001, Australia, Email imaina.widagdo@unisa.edu.au



**Purpose:** Identifying patient health conditions in observational studies is essential for accurately measuring healthcare practices and planning effective health policy interventions. This analysis evaluates the validity of the Rx-Risk Index, a tool that uses medication dispensing data to identify patient comorbidities and measure overall health. We examined an updated version of the Rx-Risk Index, reflecting changes in treatment practices, to assess its validity as a tool for identifying specific health conditions and as a measure of overall health to aid in risk adjustment in observational studies.

**Patients and Methods:** We conducted a validation study using two Australian linked health datasets, the Person-Level Integrated Data Asset (PLIDA) and the National Health Data Hub (NHDH), from 2010 to 2018, focusing on individuals aged 65 years or older. The sensitivity, specificity, PPV/NPV, Cohen's kappa, and F1 scores were used to assess agreement between Rx-Risk Index conditions and two reference standards: patient self-reported conditions and hospital diagnosis. The Rx-Risk Index's predictive validity for one-year mortality was also evaluated using logistic regression, with model fit assessed by AIC and c-statistic.

**Results:** Data were analysed from 3,959 individuals in PLIDA and 157,709 individuals in NHDH. The Rx-Risk Index showed high sensitivity (≥75%) for diabetes, chronic airways disease, hyperlipidemia, and epilepsy against both self-reported conditions and hospital diagnoses. However, hyperlipidemia and hypertension showed lower specificity (<70%). High PPVs (≥78%) were observed for diabetes and renal failure. The agreement between the Rx-Risk Index and self-reported conditions was stronger (Cohen's kappa: 0.41–0.81 for 7 conditions) than between Rx-Risk Index and ICD10-AM diagnoses (kappa: 0.73 for one condition). The Rx-Risk Index was a strong predictor of one-year mortality, with c-statistic of 0.820 (95% CI: 0.817–0.825).

**Conclusion:** Selected Rx-Risk Index conditions are reasonable proxies for identifying specific conditions, particularly those requiring pharmacological management. The Rx-Risk Index was a strong predictor of one-year mortality, suggesting it is a valid measure of overall health. This study demonstrates the Rx-Risk Index's potential to enhance disease classification and risk adjustment in observational studies, supporting informed decision-making in health policy planning.

**Plain language summary:** Administrative health claims data are increasingly accessible and play a crucial role in public health research. However, these data often lack detailed information about individual health conditions. Accurately identifying a patient's health condition, or comorbidities, is essential for ensuring healthcare safety, quality, effective policy planning, and evaluation. This study evaluated a tool called the Rx-Risk Index, which uses prescription data to identify health conditions and adjust for health risks in epidemiological analysis. We have updated the list of medicines and conditions included in this tool to ensure alignment with evolving treatment practices.

Researchers used data from two large Australian health databases for people aged 65 and older to test the Rx-Risk Index. They compared the conditions identified by this tool to patient reports and hospital records. The Rx-Risk Index was effective at identifying conditions like diabetes and chronic airways disease, which have specific medication treatments. However, it was less accurate for high blood pressure, as medications for this condition can be used to treat various other conditions. Despite this limitation, the Rx-Risk Index was a strong predictor of whether someone might die within a year.

Received: 4 October 2024
Accepted: 19 February 2025
Published: 19 March 2025

Overall, the Rx-Risk Index is a valuable tool for identifying certain health conditions, especially those requiring specific medications, and for adjusting health risks in epidemiological research. However, its accuracy varies by condition, and it may need adjustments for use in countries with different medication practices.

**Keywords:** prescription claims data, disease classification, risk adjustment, predictive validity, observational study, pharmacoepidemiology

# Introduction

Health claims data are increasingly available and have contributed to studies of healthcare safety and quality as well as health surveillance, policy planning and evaluation.[1,2] However, reliable clinical records on diagnosis information are not always available in these data sources. When available, it is often derived from hospital records, which are inherently biased toward sicker populations and only available for those that have a hospital visit. Therefore, medicines use data can serve as a valuable proxy for identifying individuals with certain diseases as many medicines are used for specific diseases. Generating comorbidity scores based on prescription medicine use can help to estimate the overall burden of health conditions, to support strategies for confounding risk adjustment.[1]

The Rx-Risk Index is an example of a tool that identifies specific conditions for which medicines can be used to treat. The premise is that patients, if using specific medicines to treat a condition, can be considered to have that condition. The Rx-Risk Index was first developed in 2003, called the Rx-Risk-V Index,[3] which determined an individual's current health condition based on their prescription medicine dispensing over a discrete period.[3] The index only includes health conditions for which a medicine could be prescribed based on conditions identified from the Chronic Disease Score (CDS).[4]

Older adults often experience multiple health conditions, which can greatly impact overall health. Improving tools that can help to measure overall health can help to more appropriately understand health outcomes in research and policy. The validity of the original Rx-Risk-V Index in identifying specific diseases was evaluated in an older Australian population, where the Index demonstrated high specificity but low to moderate sensitivity compared to self-reported conditions.[5] Additionally, when the Index was used to calculate a weighted sum of identified health conditions as an overall score, it showed a significant relationship with self-rated health ($p < 0.001$). The study also found that the Rx-Risk-V Index was strong predictor of mortality, with hazard ratios of 1.079 (95% CI 1.045–1.114).[5]

In 2018, to improve the usability of the index across different health systems, the index was mapped to the World Health Organisation's (WHO) Anatomical Therapeutic Chemical (ATC) classification system.[6] This first ATC-coded Rx-Risk Index version consisted of 46 health conditions mapped to the ATC standardised international medicines classification system by consensus between two pharmacists. While the study found that the index was highly predictive of one-year mortality in older population (c-statistic = 0.833, 95% confidence interval (CI): 0.829–0.837), it was not validated for disease identification.[7]

Due to continual advances in disease management and changes in prescribing practice, the Rx-Risk Index requires regular update and re-validation. In this study, we built upon the original ATC-coded version of the Rx-Risk Index by updating it to reflect current medicine management practices and re-validating its performance. Specifically, we aimed to 1) evaluate the accuracy of the Rx-Risk Index in identifying health conditions by determining the agreement between conditions identified using the Rx-Risk Index compared to patient self-reported conditions and hospital diagnoses and 2) assess the effectiveness of the Rx-Risk Index as a measure of overall health by determining the validity of the Rx-Risk Index Score in predicting one-year mortality.

# Materials and Methods

## Data Sources

Two Australian national linked datasets were used: Person-Level Integrated Data Asset (PLIDA) and National Integrated Health Services Information (NIHSI), which has transitioned to the National Health Data Hub (NHDH) datasets.[8,9] Both

the PLIDA and NHDH datasets provide comprehensive and nationally representative insights into health service utilisation and health outcomes across the Australian population.

PLIDA integrates datasets from health (excluding hospitalisation), education, social services, income and taxation, employment, and population surveys including the National Health Survey (https://www.abs.gov.au/about/data-services/data-integration/integrated-data/multi-agency-data-integration-project-PLIDA). The PLIDA datasets used in this analysis were the National Health Survey (NHS) 2017–18 and Pharmaceutical Benefits Scheme (PBS) prescription claims data from the 2017–18 financial year (Australian financial year is between 1 July and 30 June the following year). NHDH is a national linked health administrative dataset that includes data on hospitalisation encounters, primary healthcare visits funded under the Medicare Benefits Schedule (MBS), medicines dispensed on the national PBS, Residential Aged Care services data, and National Death Index (NDI)(https://www.aihw.gov.au/reports-data/nhdh/about). The NHDH dataset used in this analysis was the NHDH version 0.5 which was the available version at the time of analysis. NHDH 0.5 includes data from 2010–11 to 2016–17 financial years.

## Study Population

Individuals in the PLIDA dataset were eligible for inclusion if they were aged 65–100 in the 2017–18 NHS record and had a unique PLIDA linkage identifier, which enabled linkage between NHS and PBS claims data. Linkage rates were consistently high across older age groups, with 93.9% for those aged 65–74, 95.5% for those aged 75–84, and 96.5% for those aged 85 and over, ensuring comprehensive representation of the older population.[8] The Rx-Risk Index was applied to PBS claims data from the 2017–18 financial year.

For the NHDH dataset, we included individuals aged 65–100 who were alive as of 30 June 2016. To ensure comprehensive healthcare utilisation data, these individuals were required to have at least one MBS or PBS claim between 1 July 2012 and 30 June 2015 and at least one hospitalisation record from 1 July 2010 to 30 June 2016. The inclusion of at least one hospitalisation record allowed the use of diagnosis codes from admitted patient care data to validate disease diagnoses identified through PBS claims. The Rx-Risk Index was applied to PBS claims data from the 2015–16 financial year.

## Rx-Risk Index Update and Application

The updated Rx-Risk Index now includes a total of 52 conditions and incorporates new medicines for existing conditions as applicable. Nine new conditions have been added: constipation, cystic fibrosis, inflammatory autoimmune diseases, macular degeneration, multiple sclerosis, myasthenia gravis, opioid dependence, pain (paracetamol), and pain (gabapentinoid). Additionally, two conditions have been removed: malnutrition and ischemic heart disease (hypertension-specific), with ischemic heart disease (hypertension-specific) merged into the hypertension category (Table 1). This update was conducted through consensus among three pharmacists.

Rx-Risk conditions were identified if there was at least one dispensing of a medicine mapped to the Rx-Risk condition category within a 12-month period (2017–18 for PLIDA dataset and 2015–16 for NHDH dataset). Additionally, a sensitivity analysis was conducted requiring two or more dispensings of a medicine mapped to the Rx-Risk condition category within the two different time periods (3 and 12 months) to indicate the presence of that condition.

## Reference Standard

To determine the validity of the Rx-Risk Index disease classification, we compared them to self-reported conditions from the National Health Survey (NHS) in PLIDA and hospital diagnoses (ICD-10-AM codes) in NHDH. Out of the 52 Rx-Risk conditions, we were able to map 19 conditions to the self-reported conditions in PLIDA and to ICD-10 diagnosis codes in NHDH (Table 2). Hospitalisations with relevant ICD-10 codes in either primary or secondary diagnoses were used to indicate the presence of the condition.

The NHDH dataset was used for the predictive validity analysis, with death in the 2016–17 financial year as the outcome. The Rx-Risk Index, applied to PBS claims data from the 2015–16 financial year, was used as the predictor variable with the index date set within this period.

**Table 1** Updated Rx-Risk Index

| Rx-Risk Condition | Medicine/Medicine Group Name | WHO ATC Code |
|---|---|---|
| **Alcohol dependency** | Drugs used in alcohol dependence | N07BB |
| **Allergies** | Antihistamines for systemic use | R06 (excluding injection, RPBS only) |
| **Anticoagulants** | Anticoagulants | B01A (except: B01AC, B01AD) |
| **Antiepileptics** | Antiepileptics | N03A (except N03AX16) |
| **Antiplatelets** | Platelet aggregation inhibitors excl. heparin | B01AC (except: B01AC09, B01AC11, B01AC27) |
| **Antipsychotics** | Antipsychotics | N05A (exclude N05AB04 and N05AN01) |
| **Anxiety/Sedative** | Anxiolytics, hypnotics and sedatives | N05B, N05C |
| **Arrhythmia** | Digoxin, antiarrhythmics class I and III, sotalol | C01AA05, C01B, C07AA07 |
| **Benign prostatic hyperplasia (BPH)** | Drugs used in BPH | G04C (in male only) |
| **Bipolar disorder** | Lithium | N05AN01 |
| **Chronic airways disease** | Drugs for obstructive airway diseases | R03 |
| **Congestive heart failure** [a] | Ivabradine, spironolactone, eplerenone, metoprolol, bisoprolol, nebivolol, carvedilol, sacubitril and valsartan, sulfonamides plain (loop diuretics) **AND** selective beta blockers /alpha-beta blockers/angiotensin-converting enzyme inhibitors/angiotensin receptor blockers) | C01EB17, C03DA01, C03DA04, C07AB02, C07AB07, C07AB12, C07AG02, C09DX04, C03CA + (C07AB or C07AG or C09A or C09C) |
| **Constipation (NEW)** | Drugs for constipation | A06 |
| **Cystic fibrosis (NEW)** | Ivacaftor, ivacaftor+lumacaftor, ivacaftor +tezacaftor, dornase alfa | R07AX02, R07AX30, R07AX31, R05CB13 |
| **Dementia** | Anti-dementia drugs | N06D |
| **Depression** | Antidepressants | N06A (except N06AX12) |
| **Glaucoma** | Antiglaucoma preparations and miotics | S01E |
| **Gastroesophageal reflux disease (GORD)** | Drugs for GORD | A02B |
| **Gout** | Antigout preparations | M04 |
| **Diabetes** | Drugs used in diabetes | A10 |
| **Hepatitis B** | Adefovir, entecavir, tenofovir disoproxil, lamivudine | J05AF08, J05AF10[b], J05AF07[b], J05AF05 (limit to 100mg tablet only as PBS specific indication for Hepatitis B) |
| **Hepatitis C** | Antivirals for treatment of Hepatitis C viral infections, daclatasvir, sofosbuvir | J05AP, J05AP07, J05AP08 |
| **Human immunodeficiency virus (HIV)** | Protease inhibitors, antivirals for treatment of HIV infections, combinations, nucleoside and nucleotide reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, enfuvirtide, maraviroc, integrase inhibitors | J05AE, J05AR, J05AF[c], J05AG J05AX07, J05AX09 J05AJ |
| **Hypothyroidism** | Thyroid hormones | H03AA |

(*Continued*)

**Table 1** (Continued).

| Rx-Risk Condition | Medicine/Medicine Group Name | WHO ATC Code |
|---|---|---|
| **Hyperthyroidism** | Antithyroid | H03B |
| **Ischaemic heart disease (IHD): angina** | Glyceryl trinitrate, isosorbide di/mononitrate, nicorandil, perhexiline | C01DA02-C01DA14, C01DX16, C08EX02 |
| **Liver disease** | Rifaximin, ursodeoxycholic acid | A07AA11, A05AA02 |
| **Hyperlipidaemia** | Lipid modifying agents | C10 |
| **Hypertension**[f] | Antihypertensives (except medicines for pulmonary arterial hypertension), diuretics, beta blockers, calcium channel blockers, renin-angiotensin system agents | C02 (except C02KX), C03[d], C07[e], C08, C09[d] |
| **Hyperkalaemia** | Polystyrene sulfonate | V03AE01 |
| **Incontinence** | Drugs for urinary frequency and incontinence, propantheline | G04BD, A03AB05 |
| **Inflammation/pain** | Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) | M01A |
| **Inflammatory - glucocorticoids** | Glucocorticoids | H02AB |
| **Inflammatory autoimmune diseases**[g] **(NEW)** | Sulfasalazine, mesalazine, olsalazine, balsalazide, prednisolone, hydrocortisone, budesonide, vedolizumab, guselkumab, tildrakizumab, risankizumab, ustekinumab, secukinumab, ixekizumab, etanercept, infliximab, adalimumab, certolizumab, golimumab, methotrexate, tocilizumab, leflunomide, abatacept, tofacitinib, baricitinib, specific antirheumatic agents, hydroxychloroquine | A07EC01-A07EC04, A07EA01-A07EA02, A07EA06, L04AA33, L04AC16, L04AC17, L04AC18, L04AC05, L04AC10, L04AC13, L04AB01, L04AB02, L04AB04, L04AB05, L04AB06, L04AX03, L04AA07, L04AA13, L04AA24, L04AA29, L04AA37, M01C, P01BA02 |
| **Macular degeneration (NEW)** | Antineovascularisation agents | S01LA |
| **Malignancies** | Antineoplastic agents | L01 |
| **Migraine** | Antimigraine preparations | N02C |
| **Multiple sclerosis (NEW)** | Interferon beta-1a, interferon beta-1b, peginterferon beta-1a, glatiramer acetate, natalizumab, fingolimod, teriflunomide, alemtuzumab, ocrelizumab, cladribine, siponimod, dimethyl fumarate | L03AB07, L03AB08, L03AB13, L03AX13, L04AA23, L04AA27, L04AA31, L04AA34, L04AA36, L04AA40, L04AA42, L04AX07 |
| **Myasthenia gravis (NEW)** | Pyridostigmine | N07AA02 |
| **Opioid dependence (NEW)** | Drugs used in opioid dependence | N07BC |
| **Osteoporosis/Paget's** | Drugs affecting bone structure and mineralisation, Raloxifene, Teriparatide | M05B, G03XC01, H05AA02 |
| **Pain - Opioid** | Opioids | N02A |
| **Pain - Gabapentinoid (NEW)** | Gabapentinoids | N03AX16 (Pregabalin), N03AX12 (Gabapentin - RPBS only) |
| **Pain - Paracetamol (NEW)** | Paracetamol | N02BE01 |

(*Continued*)


https://doi.org/10.2147/CIA.S494145

**Table 1** (Continued).

| Rx-Risk Condition | Medicine/Medicine Group Name | WHO ATC Code |
|---|---|---|
| **Pancreatic insufficiency** | Multienzymes (lipase, protease etc.) | A09AA02 |
| **Parkinsons/Restless leg syndrome (RLS)** | Anti-Parkinson drugs | N04 |
| **Psoriasis** | Antipsoriatics | D05 |
| **Pulmonary hypertension** | Antihypertensives for pulmonary arterial hypertension, sildenafil, tadalafil, iloprost, epoprostenol, selexipag | C02KX, G04BE03, G04BE08 - excluding RPBS list, B01AC09, B01AC11, B01AC27 |
| **Renal disease** | Erythropoietin, darbepoetin alfa, methoxy polyethylene glycol-epoetin beta, sevelamer, lanthanum carbonate, sucroferric oxyhydroxide | B03XA01-B03XA03, V03AE02, V03AE03, V03AE05 |
| **Smoking cessation** | Drugs used in nicotine dependence (N07BA), bupropion | N07BA01-N07BA03, N06AX12 |
| **Transplant** | Calcineurin inhibitors (L04AD), Mycophenolic acid, sirolimus, everolimus | L04AD, L04AA06, L04AA10, L04AA18 |
| **Tuberculosis (TB)** | Hydrazides, combination treatment for tuberculosis | J04AC, J04AM |

**Notes**: (https://www.pbs.gov.au/browse/rpbs); [a]Must have at least two medicines prescribed with one of those medicines having an ATC code from C03CA01–C03CC01 and the other having an ATC code from either C09AA01–C09AX99 or C09CA01–C09CX99; PBS – Pharmaceutical Benefits Scheme (https://www.pbs.gov.au/pbs/home); [b]to use with Authority Streamlined (AS) codes: 6980,6992,6983,6984,10,362 – for hepatitis B; [c]to use with AS codes 6998 and 6982 – for HIV use, to exclude: Lamivudine 100mg tab, J05AF08 (Adefovir), J05AF10 (Entecavir)); [d]Can have medicine dispensed with an ATC code C03CA01–C03CC01 or C09AA01–C09AX99 but not both; as this would indicate chronic heart failure; [e]If PBS item code is not 8732N, 8733P, 8734Q, 8735R; [f]Excluding medicines identified as congestive heart failure or ischemic heart disease-angina. [g]Including: rheumatoid arthritis, ulcerative colitis, Crohn's disease, systemic lupus erythematosus, psoriatic arthritis, severe chronic plaque psoriasis, ankylosing spondylitis.

**Abbreviations**: WHO ATC, World Health Organization Anatomical Therapeutic Chemical classification system; RPBS, Repatriation Pharmaceutical Benefits Scheme

**Table 2** Rx-Risk Conditions and Corresponding Self-Reported Conditions as Recorded in the Person-Level Integrated Data Asset (PLIDA) Dataset and in the National Health Data Hub (NHDH) Hospital Diagnosis Code (ICD-10-AM)

| Rx-Risk Condition | PLIDA - NHS Code and Description | NHDH – Hospital Diagnosis Code (ICD-10-AM) |
|---|---|---|
| Hyperlipidaemia | 31. High cholesterol | E78 |
| Heart failure | 92. Heart failure | I50 |
| Depression | 39. Feeling depressed | F3, F4 |
| | 40. Depression | F3, F4 |
| | 41. Manic Episode | F3, F4 |
| | 42. Bipolar affective disorder | F3, F4 |
| | 43. Other mood (affective) disorders | F3, F4 |
| | 44. Feeling anxious, nervous or tense | F3, F4 |
| | 45. Anxiety disorder, including generalised anxiety disorder | F3, F4 |
| | 46. Panic Disorder | F3, F4 |
| | 47. Panic Attacks | F3, F4 |
| | 48. Phobic anxiety disorders | F3, F4 |

(*Continued*)

**Table 2** (Continued).

| Rx-Risk Condition | PLIDA - NHS Code and Description | NHDH – Hospital Diagnosis Code (ICD-10-AM) |
|---|---|---|
| | 49. Obsessive Compulsive disorder | F3, F4 |
| | 50. Post-Traumatic Stress disorder | F3, F4 |
| | 51. Other anxiety related disorder | F3, F4 |
| Hypertension | 88. Hypertensive disease | I10-I15 |
| Chronic airways disease* | 107. Bronchitis | J44, J45 |
| | 108. Emphysema | J44, J45 |
| | 109. Asthma | J44, J45 |
| Diabetes^ | 27. Type A Diabetes mellitus | E10-E14 |
| | 28. Type B Diabetes mellitus | E10-E14 |
| | 29. Type unknown Diabetes mellitus | E10-E14 |
| Osteoporosis/Paget* | 138. Osteoporosis | M80, M81, M82 |
| Gout | 127. Gout | M10 |
| Glaucoma | 68. Glaucoma | H40, H41, H42 |
| Epilepsy | 63. Epilepsy | G40 |
| Ischemic heart Disease-angina | 89. Angina | I20 |
| Migraine | 64. Migraine | G43 |
| Malignancies* | 14. Digestive organs Malignant neoplasms | C00-C96, Z511, Z510 |
| | 15. Respiratory & intrathoracic organs Malignant neoplasms | C00-C96, Z511, Z510 |
| | 16. Skin Malignant neoplasms | C00-C96, Z511, Z510 |
| | 17. Mesothelial & soft tissue Malignant neoplasms | C00-C96, Z511, Z510 |
| | 18. Breast Malignant neoplasms | C00-C96, Z511, Z510 |
| | 19. Female genital organs Malignant neoplasms | C00-C96, Z511, Z510 |
| | 20. Male genital organs Malignant neoplasms | C00-C96, Z511, Z510 |
| | 21. Other Malignant neoplasms | C00-C96, Z511, Z510 |
| | 22. Site unknown Malignant neoplasms | C00-C96, Z511, Z510 |
| Psoriasis | 124. Psoriasis | L40 |
| Macular degeneration | 69. Macular degeneration | H353 |
| Inflammatory autoimmune | 128. Arthritis - Rheumatoid | M32, M05, M06, M08, M353, M45, L40, K50, K51 |
| Renal failure* | 145. Kidney Disease | Z49, N184, N185 |
| Dementia | 34. Organic mental problems, including dementia | G30, F00-F03 |
| Bipolar disorder | 42. Bipolar affective disorder | F31 |
| Transplant | N/A | Z940-Z944, Z948, Z949 |

**Notes**: *for PLIDA-NHS, include condition status: ever told has condition, still current and long-term and not known or not ever told, but condition current and long-term; ^for PLIDA-NHS, include condition status: ever told has condition – still current and long-term/still current but not long-term, ever told has condition not current, and not known or not ever told, but condition current and long-term.

**Abbreviations**: NHS – National Health Survey; ICD-10-AM – International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification;

## Statistical Methods for Agreement

Agreement between conditions identified by the Rx-Risk Index and the self-reported conditions or hospital diagnosis codes was assessed using Cohen's kappa statistic and F1 score. Cohen's kappa measures the level of agreement between two raters (classifiers) in classifying binary categories. The F1 score is a harmonic mean of positive predictive value (precision) and sensitivity (recall); it describes the proportion of correct positive predictions. Both kappa and F1 scores range from 0 (worse) to 1 (best). The kappa value can be interpreted as follows: no agreement (≤0), none to slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect agreement (0.81–1.00).

In addition to these measures of agreement, we evaluated sensitivity and specificity as measures of the validity of the Rx-Risk Index in identifying people with and without the conditions. Sensitivity and specificity assess the ability of the Rx-Risk Index to correctly classify individuals as having or not having a condition based on medicines dispensed. We also calculated positive and negative predictive values (PPV and NPV) to provide practical information about the likelihood of a condition being present given a positive or negative Rx-Risk result. However, it is important to note that PPV and NPV are influenced by the prevalence of the condition in the study population and are not measures of the validity of the Rx-Risk Index. The values for sensitivity, specificity, and predictive values were interpreted as follows:

- Sensitivity (recall): proportion of people with the condition as defined by hospital diagnosis or self-reported condition who had medicines dispensed in the corresponding Rx-Risk condition
- Specificity: proportion of people without the condition as defined by hospital diagnosis or self-reported condition who did not have medicines dispensed in the corresponding Rx-Risk condition
- Positive predictive value (PPV/precision): proportion of people with medicines dispensed in a Rx-Risk condition who also have the condition as defined by hospital diagnosis or self-reported condition
- Negative predictive value (NPV): proportion of people without medicines dispensed in a Rx-Risk condition who also did not have the condition as defined by hospital diagnosis or self-reported condition

## Statistical Methods for Predictive Validity

A baseline logistic regression model was performed with one-year mortality as the outcome and age and sex as the predictors. The Rx-Risk condition scores and individual Rx-Risk conditions were added to the baseline model separately. In this analysis, three forms of the Rx-Risk index were generated: an unweighted score, a weighted score, and an individual condition with an indicator variable indicating the presence or absence of each of the 52 Rx-Risk categories. Cystic fibrosis and Hepatitis B were removed as categories in the calculation of the RxRisk score for predictive validity analysis as the number of individuals with these conditions was less than 10. Hence, the Rx-Risk score was calculated as the count of the number of different Rx-Risk conditions a person had been treated for with a possible score ranging from 0 to 50.

The overall goodness of fit for each model was compared with the baseline model using the Akaike information criterion (AIC).[10] The difference between the AIC values of two models must be greater than 10 for one model to be considered superior to the other. Model discrimination was compared based on c-statistic which can range from 0 to 1, with 1 indicating perfect prediction and 0.5 indicating chance predictions. A c-statistic between 0.8 and 0.9 was considered as excellent and between 0.7 and 0.8 acceptable.[11] Using 250 bootstrap samples, 95% CIs were generated for the c-statistic.[12]

The weighted Rx-Risk score was calculated by adding the Rx-Risk conditions to a binary logistic regression model with 1-year mortality as the outcome, adjusting for age and gender. From this model, each Rx-Risk condition was weighted according to its statistical significance and the magnitude of its odds ratio.[13] The weighted Rx-Risk score for an individual was then derived as the sum of the weighted Rx-Risk conditions.

To validate the weights calculated for the Rx-Risk conditions, 10-fold cross validation was used. This subset the NHDH cohort using random sampling without replacement into 10 equal folds, where each fold is 10% of the data. For cross-validation, one-fold was chosen as the testing set and the remaining 9 folds were used as the training set. The training set was used to calculate the condition weights and then used to calculate a weighted Rx-Risk score for patients

in training and testing sets. A binary logistic regression model with one-year mortality including age, sex and weighted Rx-Risk score was built separately for the training and testing sets and the c-statistic recorded. This process was repeated 10 times until each fold was used as a testing set once. This process calculated 20 c-statistics, 10 each on the training and testing set. The average c-statistic was recorded for each set. Statistical analyses were undertaken using SAS version 9.4 (SAS Institute, Cary, NC, US) and R version 4.1.0.

## Results

The PLIDA analytic cohort included 3,959 NHS participants, with an average age of 74.2 years (Standard Deviation (SD): 6.6) and 56% females. The NHDH analytic cohort consisted of 157,709 patients, with an average age of 74.6 years (SD: 7.6) and 53% females. Overall, the prevalence of each of the 19 selected conditions was higher when identified as self-reported conditions compared to hospital diagnosis codes, except for heart failure, diabetes, ischemic heart disease (angina), and dementia (Table 3). The largest discrepancy was observed in the prevalence of hyperlipidaemia, with the

**Table 3** Prevalence and Kappa Statistics of Rx-Risk Conditions Between Self-Report and Hospital Diagnosis Data

| Rx-Risk Conditions | PLIDA (N = 3,959) | | | NHDH (N = 157,709) | | |
|---|---|---|---|---|---|---|
| | Prevalence Rx-Risk FY 2017–18 (%) | Prevalence Self-Report NHS 2017–18 (%) | Kappa (95% CI) | Prevalence Rx-Risk FY 2015–16 (%) | Prevalence ICD-10-AM FY 2010–11 to 2015–16 (%) | Kappa (95% CI) |
| **Hyperlipidaemia** | 48.95 | 26.8 | 0.4016 (0.3759–0.4273) | 52.75 | 1.16 | 0.0141 (0.0131–0.0151) |
| **Heart Failure** | 9.83 | 1.69 | 0.1283 (0.0852–0.1715) | 14.91 | 6.45 | 0.3758 (0.3713–0.3802) |
| **Depression** | 20.99 | 23.06 | 0.4224 (0.3887–0.4560) | 24.58 | 5.31 | 0.1484 (0.1449–0.1518) |
| **Hypertension** | 57.16 | 43.87 | 0.4966 (0.4707–0.5224) | 59.51 | 20.06 | 0.0772 (0.0737–0.0807) |
| **Chronic airways disease** | 21.37 | 16.87 | 0.6028 (0.5711–0.6344) | 23.68 | 5.91 | 0.2671 (0.2634–0.2707) |
| **Diabetes** | 15.03 | 16.62 | 0.8133 (0.7881–0.8385) | 18.04 | 19.13 | 0.7386 (0.7337–0.7436) |
| **Osteoporosis/ Paget's** | 12.53 | 16.75 | 0.4652 (0.4268–0.5036) | 12.34 | 2.19 | 0.1179 (0.1145–0.1213) |
| **Gout** | 6.59 | 12.45 | 0.5123 (0.4678–0.5569) | 8.28 | 1.31 | 0.1532 (0.1498–0.1565) |
| **Glaucoma** | 7 | 3.59 | 0.5164 (0.4575–0.5754) | 8.05 | 0.54 | 0.0831 (0.0809–0.0855) |
| **Epilepsy** | 2.83 | 0.96 | 0.4048 (0.3056–0.5040) | 3.72 | 0.46 | 0.1783 (0.1753–0.1813) |
| **Ischemic heart disease-angina** | 5.08 | 3.86 | 0.3559 (0.2898–0.4221) | 8.35 | 5.03 | 0.2816 (0.2769–0.2864) |
| **Migraine** | 0.83 | 3.66 | 0.2596 (0.1747–0.3445) | 0.85 | 0.45 | 0.1191 (0.1144–0.1238) |
| **Malignancies** | 1.87 | 6.72 | 0.2001 (0.1417–0.2585) | 2.19 | 4.73 | 0.3242 (0.3196–0.3288) |
| **Psoriasis** | 0.88 | 3.33 | 0.1802 (0.1004–0.2599) | 1.00 | 0.17 | 0.0783 (0.0748–0.0817) |

(*Continued*)


https://doi.org/10.2147/CIA.S494145

**Table 3** (Continued).

| Rx-Risk Conditions | PLIDA (N = 3,959) | | | NHDH (N = 157,709) | | |
|---|---|---|---|---|---|---|
| | Prevalence Rx-Risk FY 2017–18 (%) | Prevalence Self-Report NHS 2017–18 (%) | Kappa (95% CI) | Prevalence Rx-Risk FY 2015–16 (%) | Prevalence ICD-10-AM FY 2010–11 to 2015–16 (%) | Kappa (95% CI) |
| **Macular degeneration** | 1.74 | 5 | 0.3579 (0.2839–0.4320) | 1.56 | 0.57 | 0.1131 (0.1086–0.1174) |
| **Inflammatory autoimmune** | 2.78 | 6.74 | 0.2683 (0.2076–0.3289) | 3.06 | 1.67 | 0.2856 (0.2809–0.2903) |
| **Renal failure** | n.p. | 3.7 | 0.1232 (0.0539–0.1926) | 0.73 | 2.13 | 0.3918 (0.3875–0.3961) |
| **Dementia** | 0.86 | 1.21 | 0.3594 (0.2254–0.4934) | 1.70 | 4.30 | 0.2664 (0.2619–0.2708) |
| **Bipolar disorder** | n.p. | 0.58 | 0.1221 (−0.0348–0.2791) | 0.29 | 0.33 | 0.2899 (0.2850–0.2948) |

**Abbreviations**: PLIDA – Person-Level Integrated Data Asset; NHDH – National Health Data Hub; NHS – National Health Survey; FY – financial year; ICD-10-AM – International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification; CI – confidence interval; n.p. – not published

Rx-Risk Index identifying approximately 50%, self-reported data showing 26.8%, and hospital diagnosis codes identifying only 1.2% (Table 3).

We found moderate to almost perfect agreement between self-reported conditions and Rx-Risk Index conditions for depression, hypertension, chronic airways disease, diabetes, osteoporosis/Paget's, gout, and glaucoma (kappa 0.51–0.81). Substantial agreement between hospital diagnosis records and Rx-Risk was evident only for diabetes (kappa 0.73) (Table 3). The sensitivity of the Rx-Risk Index for identifying heart failure was 72% when assessed against hospital diagnosis but only 52% when compared to self-report.

When using self-reported conditions as reference standard, sensitivity was greater than 75% for hyperlipidemia, hypertension, epilepsy, diabetes, glaucoma, and chronic airways disease, suggesting that of those who self-reported these conditions, a high proportion were dispensed medicines identified by the Rx-Risk Index for that condition. Renal failure, bipolar disorder, psoriasis, urinary incontinence, malignancies and migraine, however, all had low sensitivity, suggesting that of those who self-reported these conditions only a small proportion were dispensed medicines identified by the Rx-Risk Index for that condition.

Specificity was generally high, above 90%, for most conditions identified using Rx-Risk compared to self-reported conditions, suggesting that of those who did not self-report these conditions a high proportion were not dispensed medicines identified by the Rx-Risk index for that condition. The Rx-Risk Index had low specificity for identifying hyperlipidemia and heart failure (65% and 54% respectively) despite their high sensitivity. This suggests that a high proportion of patients who do not self-report having these conditions are taking medicines for hyperlipidemia and heart failure.

The Rx-Risk Index conditions for diabetes, gout, and renal failure had high PPV (>80%) (Figure 1A), indicating that most patients receiving medications for these conditions also self-reported having them. Migraine and macular degeneration had reasonable PPV (>70%). NPV was high across all conditions (>80%) (Figure 1B), suggesting that over 80% of patients without Rx-Risk identified condition-specific medicines did not self-report having the condition. Conversely, approximately 20% of patients who reported having a condition did not have any record of condition-specific medicines dispensed. This discrepancy could be due to several factors, including non-pharmacological management, the use of over-the-counter medicines or medicines not subsidised under the PBS, alternatively this could be due to potential errors in self-reporting.

When using hospital diagnoses as reference standard, the Rx-Risk identified conditions had on average low PPV, with only diabetes and renal failure showing high positive predictive values of around 80% (Figure 1C). Negative predictive value of the Rx-Risk Index was high for all conditions (Figure 1D).

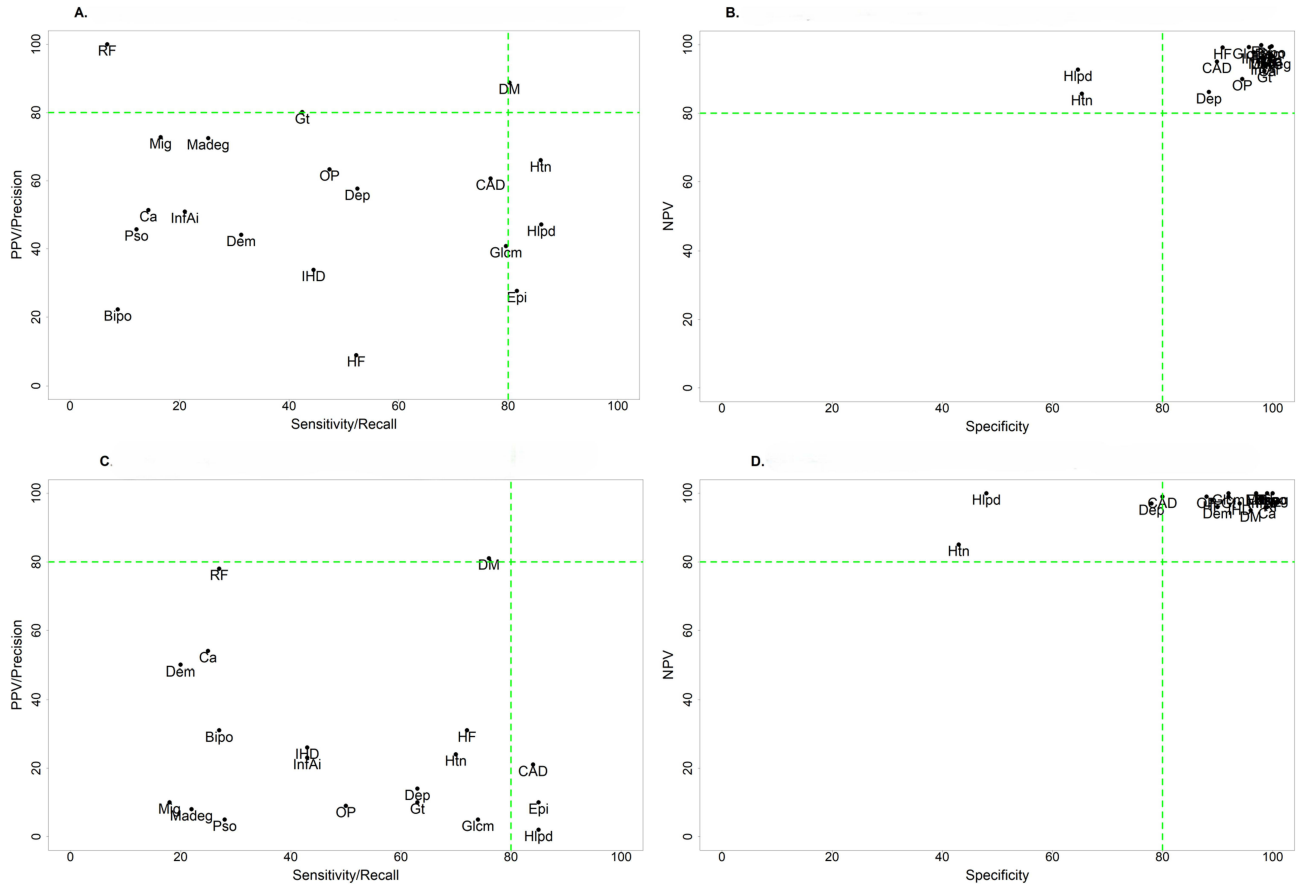


**Figure 1** Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of Rx-Risk Index against self-reported conditions and ICD10-AM. (**A**) Sensitivity and PPV of the Rx-Risk Index compared to self-reported conditions (**B**). Specificity and NPV of the Rx-Risk Index compared to self-reported conditions (**C**). Sensitivity and PPV of the Rx-Risk Index compared to ICD10-AM codes (**D**). Specificity and NPV of the Rx-Risk Index compared to ICD10-AM codes.

The sensitivity analysis for disease identification agreement found that overall, the prevalence of a specific condition was inversely related to the number of dispensings required and the time window duration to define a condition (Supplementary Tables 1 and 2).

The baseline model, with only age and sex only, predicted one-year mortality moderately well (c-statistic = 0.780). The addition of the unweighted Rx-Risk score to the model increased the performance of the model (c-statistic = 0.820). The addition of indicator variables for each conditions performed best (c-statistic 0.856) (Table 4). Results of the 10-fold

**Table 4** Comparison of the Different Rx-Risk Scoring and Modelling Methods to Predict 1-year Mortality

| | AIC* | Difference in AIC [1] | C-statistic | 95% CI |
|---|---|---|---|---|
| Primary analysis (1 or more dispensing) | | | | |
| Base model (BM): age and sex | 97,809 | - | 0.780 | 0.776–0.784 |
| BM + unweighted Rx-Risk | 93,590 | 4,219 | 0.820 | 0.817–0.825 |
| BM + weighted Rx-Risk | 87,973 | 9,863 | 0.812 | 0.808–0.815 |
| BM + 50 Rx-Risk conditions | 86,848 | 10,961 | 0.856 | 0.853–0.859 |

(*Continued*)


https://doi.org/10.2147/CIA.S494145

**Table 4** (Continued).

| | AIC* | Difference in AIC [1] | C-statistic | 95% CI |
|---|---|---|---|---|
| Sensitivity analysis (2 or more dispensings) | | | | |
| BM + unweighted Rx-Risk | 93,974 | 3,835 | 0.817 | 0.813–0.821 |
| BM + weighted Rx-Risk | 88,370 | 9,439 | 0.855 | 0.852–0.859 |
| BM + 50 comorbidity indicators | 87,337 | 10,472 | 0.860 | 0.857–0.863 |

**Notes**: *The model with the lowest AIC value is considered the best fit; [1]AIC score compared with the AIC score of the base model. A model with a lower AIC score of 10 (or more) is considered superior.
**Abbreviation**: AIC – Akaike Information Criteria; CI – confidence interval.

cross validation were consistent with the primary analysis. The average c-statistic on the 10 training datasets was 0.857 and for testing datasets was 0.856.

In the sensitivity analysis, where two or more prescriptions dispensed in a 12-month period were required to indicate the presence of a condition, the c-statistics were similar to the main analysis. The unweighted Rx-Risk c-statistic was 0.817, the weighted Rx-Risk c-statistic was 0.855, and individual conditions c-statistic was 0.860. Despite differences in the weights assigned to some Rx-Risk conditions compared to the primary analysis, which required only one or more prescriptions, the overall performance remained consistent (Supplementary Table 3).

## Discussion

In this study, we found that conditions identified by the updated Rx-Risk Index had reasonable agreement with self-reported conditions but poor agreement with hospitalisation records for nearly all conditions except diabetes and renal failure. This variation likely reflects the nature of certain conditions, such as those primarily managed in outpatient settings or associated with preventive care (eg, hyperlipidaemia), which are less likely to be documented in hospital records. Additionally, with self-reported conditions, some individuals may be taking medications for conditions they do not report, either due to low health literacy or lack of awareness of the condition for which the medication is prescribed.

A previous study assessed the sensitivity and specificity of the Rx-Risk-V Index (non-ATC coded) conditions and self-reported conditions using the Australian Longitudinal Study on Ageing (ALSA) survey and found that the Rx-Risk category for thyroid disease was the only condition with high sensitivity (82%).[5] The differences in sensitivities could be attributed to the use of self-reported medicines in the ALSA study, which is susceptible to recall bias or underreporting, compared to the pharmacy claims data used in the present study, which provides a more comprehensive and objective record of dispensed medications.

In the Irish LongituDinal study on Ageing (TILDA), which used a modified first ATC-coded Rx-Risk Index,[7] the Rx-Risk Index demonstrated moderate-to-high agreement with self-reported conditions for asthma, hyperlipidaemia, glaucoma, and diabetes (kappa 0.41–0.81).[14] In our study, the Rx-Risk Index showed moderate-to-high agreement with self-reported conditions for seven conditions – depression, hypertension, chronic airways disease, diabetes, osteoporosis/Paget's disease, gout, and glaucoma demonstrated moderate-to-high agreement with self-reported conditions (kappa 0.41–0.81).

The PPV of the Rx-Risk Index was high for conditions like renal failure, diabetes, gout, migraine, and macular degeneration, indicating Rx-Risk Index effectively identified these diseases, as many people with condition-specific medicines self-reported having the condition. However, for conditions with overlapping treatment options, the Rx-Risk Index may have limited utility for disease identification, as it may be unable to distinguish between conditions treated with similar medications.

Our sensitivity analysis showed that for some conditions increasing the number of dispensing required to identify a Rx-Risk Index condition or reducing the time window over which dispensings were considered increased the PPV, however there was a trade off with decreased NPV (Supplementary Table 2 and 3). This means that increased ability to correctly identify conditions comes at the cost of increased false negative results. High PPVs were harder to achieve

when hospital diagnosis was used as the reference standard as only a fraction of people will have a record of hospitalisation and not all conditions are recorded in the hospital admission records.

Epidemiological studies often aim to understand the effects of an exposure—such as a medication, environmental factor, or behavioral risk factor— on a defined outcome. However, these associations can be biased due to selective prescribing of exposures to those with existing comorbidities.[15] In this context, validated tools like the Rx-Risk Index, which identify individuals with diseases and comorbidities, can play a critical role in controlling for bias by providing a proxy for underlying baseline risk. By providing a standardized approach to disease identification based on medicines use data, the Rx-Risk Index can help as a measure of overall health, which can be used to reduce systematic error, due to confounding.[15–17] In this study, we found that the Rx-Risk score is a valid tool for adjusting for confounding. The Rx-Risk score is a strong predictor of one-year mortality, with a c-statistic of 0.8, consistent with previous study.[7] These results also align with the validation analysis of the first ATC-coded Rx-Risk Index in predicting one-year mortality (c-statistic of 0.8) in the Norwegian Epidemiologic Osteoporosis Studies (NOREPOS), which used pharmacy claims data.[18] Additionally, our current findings are consistent with those of another medicines-based comorbidity index, the Modified Chronic Disease Score (M-CDS), developed and validated in the Italian population.[19] The M-CDS was found to be predictive of one-year mortality and outperformed the Charlson Comorbidity Index (CCI), which is based on hospital diagnoses information (c-statistic: 0.761 for M-CDS vs 0.696 for CCI).[19]

## Limitations

The findings are primarily applicable to older Australians aged 65 years and older, as the study was based on Australian datasets. These data sources may not fully capture all health conditions or represent populations outside this demographic. Further validation would be required to ensure the accuracy and applicability of the Rx-Risk Index if applied to other populations, such as younger individuals or those in different healthcare systems.

The validity analysis of the updated Rx-Risk Index is constrained by the quality and scope of the reference standard. Self-reported conditions suffer from underreporting and missing information due to recall bias and health literacy, while hospital records have limited population coverage and varied disease coding rules. Consequently, despite low kappa statistics due to reference standard limitations, the Rx-Risk Index may remain valid for disease classification, especially when the medicines identified by the index are indeed used for treating a specific disease.

The Rx-Risk Index was developed as a risk adjustment measure; therefore, a specific medicine should only be mapped to one condition. This is done to avoid double counting conditions in the overall comorbidity score; however, in practice, one medicine may be indicated for use in several different conditions. In this situation, using Rx-Risk Index for disease classification can result in high sensitivity but low PPV.

## Conclusion

The use of the Rx-Risk Index as a tool for disease classification when no diagnosis information is available may be valid, particularly when specific medicines are used in disease management. The Rx-Risk Index category for conditions, such as diabetes and renal failure, were found to have high PPVs, meaning that individuals flagged based on prescription data are highly likely to have the condition. However, the variability in treatment patterns and the overlap in medications used for different conditions could limit the Rx-Risk Index's use as a standalone tool for comprehensive comorbidity measurement. To address this limitation, incorporating additional data, such as diagnostic information from primary care, could help to improve disease identification performance. Solely relying on hospital diagnoses may bias the tool to individuals with more severe conditions, potentially under estimating those with less severe cases. The updated Rx-Risk Index when used as an overall number of conditions present remains a valid comorbidity risk-adjustment measure that is strongly predictive of one-year mortality and can be an important proxy measure of overall health.

## Data Sharing Statement

The data used in this study include the Person Level Integrated Data Asset (PLIDA) and the National Health Data Hub (NHDH). Data from PLIDA are available for approved projects: https://www.abs.gov.au/about/data-services/data-integration/access-and-services and require approval by the Australian Bureau of Statistics (ABS). Data from NHDH

are available for approved projects and require approval by the Australian Institute of Health and Welfare (AIHW): https://www.aihw.gov.au/reports-data/nhdh/about.

## Ethics Approval

This analysis was undertaken as part of the Healthy Mind, Healthy Body project. The Healthy Mind, Healthy Body project has ethics approval from the Australian Institute of Health and Welfare Ethics Committee (EO2019/3/1075).

## Acknowledgments

The authors would like to acknowledge the assistance of the Australian Bureau of Statistics (ABS) and the Australian Institute of Health and Welfare (AIHW) in facilitating data provision as well as access to data analytic platform. Dr Ty Stanford for his assistance in translating our SAS macro into R package. Professor Libby Roughead for her review of the updated Rx-Risk disease classification mapping.

## Funding

This work was supported by the Australian Government Department of Health and Aged Care.

## Disclosure

Associate Professor Lisa Kalisch Ellett reports grants from Australian Government Department of Health and Aged Care, during the conduct of the study. The authors report no conflicts of interest in this work.

## References

1. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. 2014;68(3):283–287. doi:10.1136/jech-2013-202744
2. Mazzali C, Duca P. Use of administrative data in healthcare research. *Intern Emerg Med*. 2015;10:517–524. doi:10.1007/s11739-015-1213-9
3. Sloan KL, Weiner JP, Lydick E, et al. Construction and characteristics of the RxRisk-V: a VA-adapted pharmacy-based case-mix instrument. *Med Care*. 2003;41(7):761–774. doi:10.1097/01.MLR.0000064641.84967.B7
4. Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol*. 1992;45(2):197–203. doi:10.1016/0895-4356(92)90016-G
5. Vitry AI, Roughead EE, Larkin M, McDermott RA, Williams S, Georgiou A. Validity of medication-based co-morbidity indices in the Australian elderly population. *Aust N Z J Public Health*. 2009;33(2):126–130. doi:10.1111/j.1753-6405.2009.00357.x
6. World Health Organization. Guidelines for ATC classification and DDD assignment. 2023. Available from: https://www.whocc.no/atc_ddd_index. Accessed March 7, 2025.
7. Pratt NL, Kerr M, Jorm LR, Roughead EE. The validity of the rx-risk comorbidity index using medicines mapped to the anatomical therapeutic chemical (ATC) classification system. *BMJ Open*. 2018;8(4):e021122. doi:10.1136/bmjopen-2017-021122
8. Australian Bureau of Statistics. Integration of the 2017-18 national health survey and the personal linkage spine. Available from: https://www.abs.gov.au/articles/integration-2017-18-national-health-survey-and-personal-linkage-spine. Accessed January 16, 2025.
9. Australian Institute of Health. National Health Data Hub. 2016-17. Available from: https://www.aihw.gov.au. Accessed March 07, 2025.
10. Bozdogan H. Akaike's information criterion and recent developments in information complexity. *J Math Psychol*. 2000;44(1):62–91. doi:10.1006/jmps.1999.1277
11. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York, NY: John Wiley & Sons, Inc.; 2000.
12. Lu CY, Roughead EE, Vitry AI, McDermott RA, Williams S. Charlson and Rx-risk comorbidity indices were predictive of mortality in the Australian health care setting. *J Clin Epidemiol*. 2011;64(2):223–228. doi:10.1016/j.jclinepi.2010.02.015
13. Johnson ML, Weiner JP, Lydick E, Fihn SD, Richardson P, Abraham NS. Adapting the Rx-Risk-V for mortality prediction in outpatient populations. *Med Care*. 2006;44(8):793–797. doi:10.1097/01.mlr.0000218804.41758.ef
14. Mannion C, McCormack J, Cahill R, O'Malley T, Bennett K. Agreement between self-reported morbidity and pharmacy claims data for prescribed medications in an older community-based population. *BMC Geriatr*. 2020;20(1):1–14. doi:10.1186/s12877-020-01684-8
15. Schneeweiss S, Maclure M. Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol*. 2000;29(5):891–898. doi:10.1093/ije/29.5.891
16. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. *Plast Reconstr Surg*. 2010;126(2):619–625. doi:10.1097/PRS.0b013e3181de24bc
17. Rivas-Ruiz F, Pérez-Vicente S, González-Ramírez A. Bias in clinical epidemiological study designs. *Allergol Immunopathol*. 2013;41(1):54–59. doi:10.1016/j.aller.2012.04.005
18. Holvik K, Årøen A, Håheim L, Schirmer H. Contribution of an extensive medication-based comorbidity index (Rx-Risk) in explaining the excess mortality after Hip fracture in older Norwegians: a NOREPOS cohort study. *BMJ Open*. 2022;12(5):e057823. doi:10.1136/bmjopen-2021-057823
19. Iommi M, Berardi A, Ricciardi W, De Luca G, Romano R. Modified-Chronic Disease Score (M-CDS): predicting the individual risk of death using drug prescriptions. *PLoS One*. 2020;15(10):e0240899. doi:10.1371/journal.pone.0240899