

Validating health conditions in a clinical registry using administrative data algorithms

Lix, L¹, Zhang, L¹, Yan, L¹, Sajobi, T², Sawatzky, R³, and Bohm, E¹¹University of Manitoba²University of Calgary³Trinity Western University

Introduction

Clinical registries are a potentially valuable resource to study the effects of medical interventions on outcomes, particularly patient-reported outcomes like health-related quality of life, which are not included in administrative data. However, because clinical registries are primarily intended for patient management and not for research, their validity must be established.

Objectives and Approach

Our objective was to validate patient self-reported health conditions in a clinical registry. Study data were from a population-based regional joint replacement registry in the Canadian province of Manitoba. The clinical registry data were linked to administrative health data. Validated administrative data algorithms for 12 conditions were defined using diagnosis codes in hospital and physician records and medication codes in prescription drug records for the period up to three years prior to the joint replacement surgery. Accuracy of the clinical registry data was estimated using Cohen's kappa coefficient, sensitivity, specificity, and positive and negative predictive values (PPV; NPV); 95% confidence intervals were also estimated. Analyses were stratified by joint type, age group, and sex.

Results

The study cohort included 20,592 individuals (average age 66.3 years; 58.4% female); 8,424 (40.9%) had a total hip replacement. Sensitivity of the clinical registry data ranged from 16% (anemia) to more than 70% (diabetes, high blood pressure, rheumatoid arthritis); specificity was greater than 92% for all conditions, except back pain and high blood pressure. PPV ranged from 19% (back pain) to 83% (diabetes). Chance-adjusted agreement was very good for diabetes (kappa: 0.74), moderate for heart disease and high blood pressure (kappa range: 0.41-0.53) and poor or fair for back pain, anemia, cancer, depression, kidney disease, liver disease, rheumatoid

arthritis and stomach ulcers (kappa range: 0.14-0.37). Estimates varied by sex (i.e., generally higher agreement for females) and age (i.e., generally lower agreement for older age groups), but not joint type.

Conclusion/Implications

Self-reported health conditions in registry data had good validity for conditions with clear diagnostic criteria, but low validity for conditions that are difficult to diagnose or rare (e.g., cancer). Linked registry and administrative data is strongly recommended to ensure valid and accurate comorbidity measures when developing risk prediction models and conducting inter-jurisdictional comparisons of patient-reported outcome measures.

