A Comprehensive Assessment of Claims-Based Algorithms for Identifying Individuals with Diabetes

Changchun Wang MS1, David Gilbertson PhD1, Jiannong Liu PhD1, Cheryl Arko BA1, Shu-cheng Chen MS1, Marshall McBean MD MSc1,2, Allan Collins MD FACP1,2
1United States Renal Data System, Minneapolis Medical Research Foundation, 2University of Minnesota Twin Cities

Introduction

- Hebert, et al. assessed a number of claims-based algorithms for identifying individuals with diabetes using Medicare Parts A and B claims (Am Jour of Medical Quality 146:270-277, 1999). Using the latest Medicare Current Beneficiary Survey (MCBS) 2001-2002 Access to Care data, we analyzed claims-based algorithms based on virtually all possible combinations of diabetic claims in Inpatient hospital (IP), Skilled nursing facility (SNF), Home health (HH), Outpatient (OP) and Physician/supplier (PB) claim files. Using our weighting system, each algorithm was assessed by different combinations of "accuracy" characteristics (sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV)) (Table 1).

- Our objective was to find the most robust algorithms among the 299 proposed algorithms used to define claim-based diabetes.

Methods

- In order to ensure that two years of claims data were available to search for evidence of diabetes, selection criteria required that patients were:
  - at least 65 years of age at the start of the time period and resided in the 50 States or the District of Columbia,
  - were eligible for Medicare Parts A and B and never enrolled in an HMO during the two-year time period.
- The MCBS Access to Care file contains self-reported information on whether or not individuals have diabetes - this information was used as our "gold standard.
- 299 claims-based algorithms were proposed. Only "OR" relationships were considered among claim sources. Sector I: 3*1 = 242 (5 sources: IP, SNF, HH, OP and PB, 3 categories for each source: 0, 1, ≥ 2)

Results

- Five common claims-based algorithms were singled out based on the top 58 algorithms in each measurement (Table 4).
- This study shows that the original methods proposed by Hebert et al. are nearly as accurate as the best algorithms found by our methodology.

Conclusions

- Based on statistic characteristics, selected claims-based algorithms could reflect patients' real status of diabetes.
- The methodology used is a systematic method that could be used to identify the best algorithms to validate any claims-based comorbid conditions where a "gold standard" is available for comparison.

Limitations

- Only "OR" situations are considered among 5 claim sources. Some "AND" relations should be explored in the future study.
- We assume self-reported diabetes status as the "gold standard."