Identifying Type 1 and Type 2 Diabetic Cases Using Administrative Data: A Tree-Structured Model

Weihsuan Lo-Ciganic, M.S.C.P., M.S.,[1] Janice C. Zgibor, R.Ph., Ph.D.,[1]
Kristine Ruppert, R.N., Dr.P.H.,[2] Vincent C. Arena, Ph.D.,[3] and Roslyn A. Stone, Ph.D.[3]

Abstract

Background:

To date, few administrative diabetes mellitus (DM) registries have distinguished type 1 diabetes mellitus (T1DM) from type 2 diabetes mellitus (T2DM).

Objective:

Using a classification tree model, a prediction rule was developed to distinguish T1DM from T2DM in a large administrative database.

Methods:

The Medical Archival Retrieval System at the University of Pittsburgh Medical Center included administrative and clinical data from January 1, 2000, through September 30, 2009, for 209,647 DM patients aged ≥18 years. Probable cases (8,173 T1DM and 125,111 T2DM) were identified by applying clinical criteria to administrative data. Nonparametric classification tree models were fit using TIBCO Spotfire S+ 8.1 (TIBCO Software), with model size based on 10-fold cross validation. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of T1DM were estimated.

Results:

The main predictors that distinguished T1DM from T2DM are age <40 years; International Classification of Disease, 9th revision, codes of T1DM or T2DM diagnosis; inpatient oral hypoglycemic agent use; inpatient insulin use; and episode(s) of diabetic ketoacidosis diagnosis. Compared with a complex clinical algorithm, the tree-structured model to predict T1DM had 92.8% sensitivity, 99.3% specificity, 89.5% PPV, and 99.5% NPV.

Abstract cont.

Conclusion:

The preliminary predictive rule appears to be promising. Being able to distinguish between DM subtypes in administrative databases will allow large-scale subtype-specific analyses of medical care costs, morbidity, and mortality.

J Diabetes Sci Technol 2011;5(3):486-493