Erratum

Erratum to “Data mining a diabetic data warehouse”

Joseph L. Breault a,b,* , Colin R. Goodall c,d, Peter J. Fos e,b

Family Medicine, Ochsner Clinic Foundation, New Orleans, LA 70121, USA
b Health Systems Management, Tulane University, New Orleans, LA 70112, USA
c AT&T Shannon Research and Technology Laboratory, Middletown, NJ 07748, USA
d Adjunct Appointment, Biostatistics, Tulane University, New Orleans, LA 70112, USA
e School of Dentistry, University of Nevada Las Vegas, Las Vegas, NV 89154, USA

Received 12 November 2002

The Publisher regrets that the Abstract of the above article contained an error. The correct version appears below.

Abstract

Diabetes is a major health problem in the United States. There is a long history of diabetic registries and databases with systematically collected patient information. We examine one such diabetic data warehouse, showing a method of applying data mining techniques, and some of the data issues, analysis problems, and results. The diabetic data warehouse is from a large integrated health care system in the New Orleans area with 30,383 diabetic patients.

Methods for translating a complex relational database with time series and sequencing information to a flat file suitable for data mining are challenging. We discuss two variables in detail, a comorbidity index and the HgbA1c, a measure of glycemic control related to outcomes. We used the classification tree approach in Classification and Regression Trees (CART®) with a binary target variable of HgbA1c >9.5 and 10 predictors: age, sex, emergency department visits, office visits, comorbidity index, dyslipidemia, hypertension, cardiovascular disease, retinopathy and end-stage renal disease.

Unexpectedly, the most important variable associated with bad glycemic control is younger age, not the comorbidity index or whether patients have related diseases. If we want to target diabetics with bad HgbA1c values, the odds of finding them is 3.2 times as high in those <65 years of age than those older. Data mining can discover novel associations that are useful to clinicians and administrators. © 2002 Elsevier Science B.V. All rights reserved.

* Corresponding author. Tel.: +1-504-243-6021; fax: +1-504-243-6052.
E-mail addresses: joebreault@tulanealumni.net (J.L. Breault), cgoodall@att.com (C.R. Goodall), peter.fos@ccmail.nevada.edu (P.J. Fos).