Data mining and surgical site infections screening

Diane Pivot[1-2],; Arnaud Florentin[2]; Christian Rabaud[3]; Laurent Bresler[4]; Serge Aho[1]; Nicolas Jay[5]

[1] Epidemiology and Hygiene Department, Dijon University Hospital, France; [2] Epidemiology and Hygiene Department, Nancy University Hospital, France; [3] Division of Infectious Diseases, University Hospital of Nancy-Brabois, Vandoeuvre-les-Nancy, France; [4] Department of Digestive Surgery, University Hospital of Nancy-Brabois, Vandoeuvre-les-Nancy, France; [5] Department of Biostatistics, University Hospital of Nancy-Brabois, Vandoeuvre-les-Nancy, France

Context and Purpose

- Surgical site infections (SSI) ➔ 3rd most commun hospital acquired infections in France

- SSI surveillance :
Complex and time consuming…BUT essential !

➔Facilitate and alleviate surveillance

The use of data mining (DM) which operates medical and administrative data as unique source ➔ interesting

A pilot study was performed

To evaluate the possibility to develop an automated detection of SSI in a university hospital

Methods

• Pilot study:

- Retrospectively : medical and administrative data (DRG)
- September and november 2014

- Interventions of digestive surgery and neurosurgery
➔ incuded

• Three methods of DM ➔ compared to assess their feasibility to retrieve relevant criteria of suspected SSI in a single data source, Diagnosis Related Group (DRG)

① Association rules methods
② Decision trees method
③ Classification rules method

Gold Standard
Diagnosis of SSI by surgeons

• DM :
➔ Set of exploratory techniques that can extract knowledge that is hidden in large databases ⇔ BIG DATA

Results

-278 interventions➔ 21 SSI confirmed ⇔ prevalence of 7,5 %

- 807 variables : 14 related SSI's preoperative risk factors and 793 related to individual care pathway

Algorithms of Decision trees and Classification rules	
Advantages	**Limits**
➔Easy to interpret ➔ Very good overall senstivity (> 90 %) ➔ Areas under the curve ROC: from 57,9 % to 70,8 %	➔ Overfitting ➔ Sensitivity concerning non-occurrence of SSI >> to detect SSI (99 % vs 33 %)

Associations rules algorithms	
Advantages	**Limits**
➔ Several criteria to evaluate the rules ➔ More relevant to detect SSI BOTH ➔**Already known risks factors for SSI:** older age; carcinologic or emergency surgery.. AND ➔ **New predictors** founded (table I)	➔More complex interpretation

Table I. Example of three new predictors derived from ICD-10 obtained

ICD-10 codes	Entitled code
T81.4	Infection following a procedure for diagnostic and therapeutic prurposes
Y83.3	Surgical operation with external anastomosis causing abnormal reaction of patient or later complications
Y83.8	Other surgeries causing abnormal patient reactions or subsequent complications

Discussion

➢ First study to use a DM strategy

➢ Use of medical administrative data from the DRG in DM ⇔ relevant method to detect SSI

➢ Organization between services would be much more efficient and SSI surveillance much improved

diane.pivot@chu-dijon.fr