



DÉCLARATION DE LIENS D'INTÉRÊTS

J'ai un ou plusieurs liens(s) d'intérêt(s) :

Fonds de recherche : ANIOS, AQUA-TOOLS, AQUAFREE



DHREAS



DÉTECTION DES INFECTIONS DU SITE OPÉRATOIRE PAR DATAMINING

Apport des données médicales issues du PMSI sur la performance
d'un modèle de détection



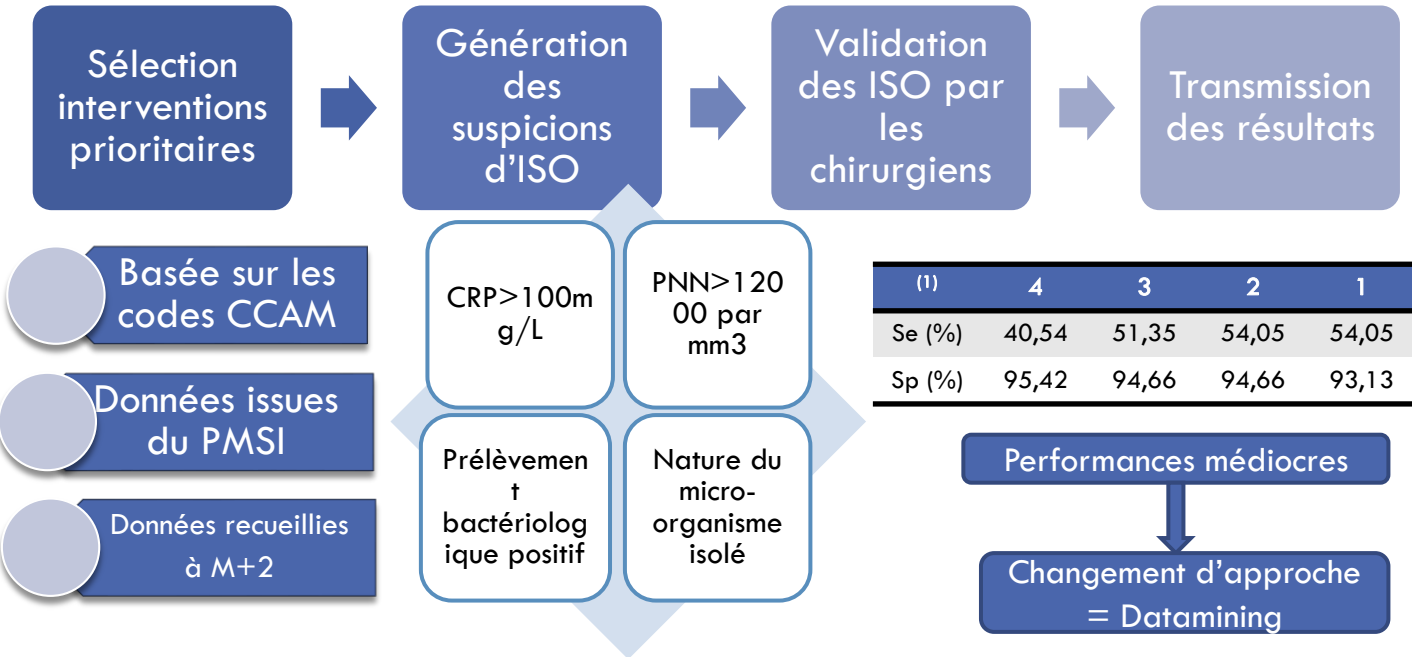
08 juin 2017

Julie LIZON PharmD, MSc, Nicolas JAY, MD, PhD, Boris MARCHAND,
Arnaud FLORENTIN, MD, MSc

- Infection du site opératoire (ISO)
 - Infection survenant dans les **30 jours suivant l'intervention, ou dans l'année si il y a eu mise en place d'un implant, d'une prothèse ou d'un matériel prothétique**
 - **3^{ème} infections nosocomiales (IN) en France⁽¹⁾ avec une incidence de moins de 1% à plus de 15% selon type de chirurgie⁽²⁾**
 - **Conséquences** : séquelles, allongement durée hospitalisation, coût important, décès
 - **Prévention** : maîtrise du risque infectieux, surveillance des ISO

(1) Institut national de Veille Sanitaire (InVS). Enquête nationale de prévalence des infections nosocomiales et des traitements anti-infectieux en établissements de santé, France. 2012.

(2) Institut national de Veille Sanitaire (InVS). Surveillance des infections du site opératoire France, 2012. Résultats. 2012.



(1) Hautemanière A, Florentin A, Hunter PR, Bresler L, Hartemann P. Screening for surgical nosocomial infections by crossing databases. J Infect Public Health. 2013 Apr;6(2):89–97.

Objectif principal

5

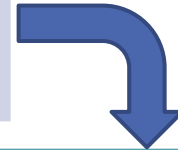
- Evaluer la contribution
 - Des données médicales issues du PMSI*
 - Diagnostics principaux (DP), Diagnostics Associées Significatifs (DAS), Codes CCAM
 - Et des modèles de prédiction de datamining pour détecter les ISO

* PMSI : Programme de Médicalisation des Systèmes d'Information

Matériel et méthode

Population

Tous les adultes bénéficiant d'une intervention cible entre le 1er janvier et le 31 août 2015
10 services de chirurgie
Durée de suivi : 30 jours post-opératoire



Critère de jugement

ISO
Gold standard : diagnostic chirurgical



Sources des données

Données microbiologiques
Données biologiques
Données médico-administratives

Matériel et méthode

7

Datamining (DM)

2497 variables dont 680 codes CCAM et 1760 codes CIM-10

Méthodes d'apprentissages supervisés

Découpage échantillon : apprentissage (70%) et validation (30%)

Modèle aléatoire

- Forêt aléatoire⁽¹⁾
 - Ajout d'une randomisation
 - Hasard dans choix des variables qui interviennent dans la construction des modèles
 - 500 arbres de décision

Modèle adaptatif

- AdaBoost⁽²⁾
 - Adaptation des modèles construits en donnant plus de poids aux observations mal ajustées ou mal prédites
 - 50 itérations

(1) Breiman L. Random Forests. Mach Learn. 45(1):5-32

(2) Freund Y, Schapire RE. Experiments with a new boosting algorithm. Mach Learn. 1996;148-56.

Matériel et méthode

Evaluation des performances

Pour tous les modèles

- Se, Sp, VPP, VPN, Courbe ROC, AUC
- Indice NRI^{*(1)}
- Calcul de l'erreur globale de classification

* Net reclassification improvement

Pour les modèles DM

- Vérification pertinence construction modèle

⁽¹⁾ Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008 Jan 30;27(2):157-72; discussion 207-12.

Résultats

□ Population

▣ 3900 patients inclus


- Moyenne âge : 56 ans +/- 19,79 [18-105]
- Sexe ratio : 0,78

	Nombre d'interventions	Détection ISO	ISO diagnostiquées	Taux d'ISO (%)
Chirurgie thoracique	118	24	16	13,54
Chirurgie gynécologique et obstétrique	639	4	3	0,47
Totalité	3900	405	192	4,92

Résultats

10

□ Performance modèles

	Algorithme actuel	Forêt aléatoire	AdaBoost
Sensibilité (%)	67,8	81,4	79,7
Spécificité (%)	92,2	97,8	97,7
VPP (%)	31,5	66,7	67,1
VPN (%)	98,1	99,0	98,9
AUC	0,827	0,988	0,979
 NRI	∅	0,212	0,176
Erreur classification (%)	9	3	3

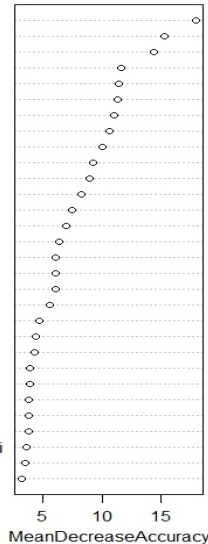
$p < 10^{-5}$

$p < 10$ $p = 0,43$

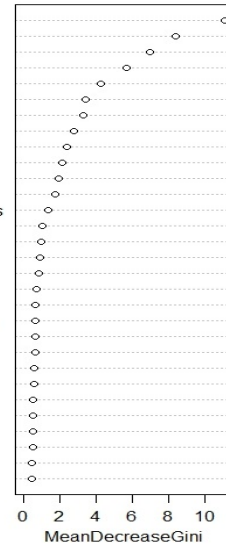
□ Pertinences variables : modèle Forêt aléatoire

- T814 (Complication infectieuse d'actes de diagnostics ou thérapeutiques)
- L022 (Présence d'un abcès cutané)
- T8138 (Désunion de la plaie opératoire)
- QZJA011 (Evacuation collection profonde)

T814
L022
score_global
score_germes
Staphylococcus.aureus
score_plvt
Escherichia.coli
score_crp
Enterococcus.faecalis
T8138
Staphylococcus.epidermidis
Klebsiella.pneumoniae
Pseudomonas.aeruginosa
duree_hospit.5j
QZJA011
Morganella.morganii
L088
T81
N823
J188
score_pnn
HHQE005
ZC-JA004
Clostridium.perfringens
J985
HZHE002
XOQK023
Corynebacterium.kroppenstedtii
Y834
var_coelio

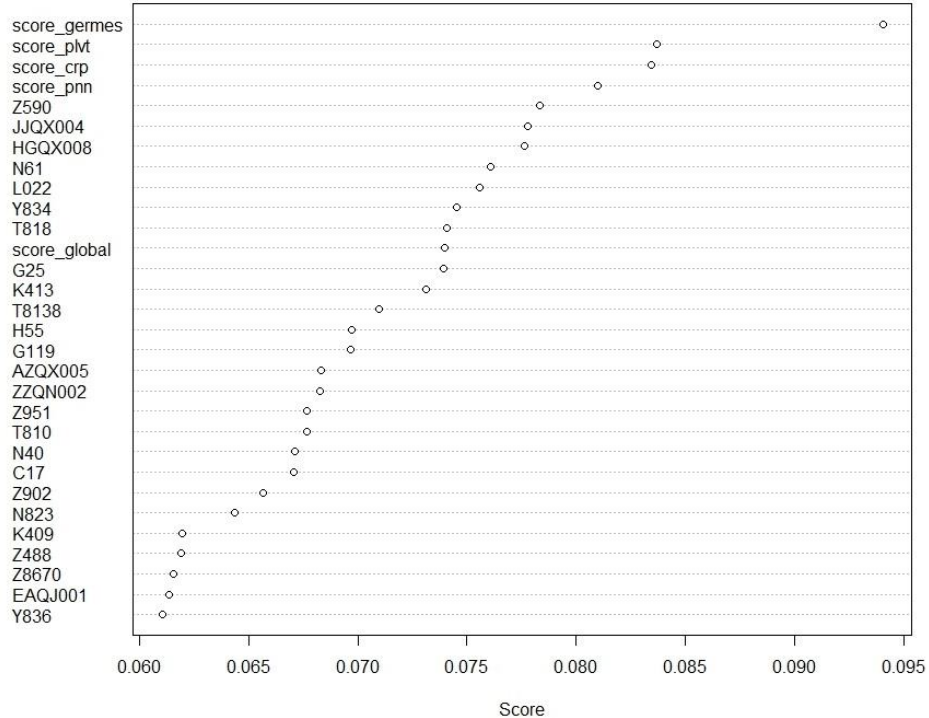


T814
score_global
score_germes
score_plvt
score_crp
L022
T81
score_pnn
T8138
Escherichia.coli
Staphylococcus.aureus
Enterococcus.faecalis
Staphylococcus.epidermidis
K650
Y832
Klebsiella.pneumoniae
R65
B952
duree_hospit.5j
Pseudomonas.aeruginosa
U82
QZJA011
Y836
ZBQK002
T818
K913
ZZQP001
J18
L088
B957



□ Pertinences variables : modèle Adaboost

- L022 (Présence d'un abcès cutané)
- T8138 (Désunion de la plaie opératoire)
- T818 (Autre complication d'actes de diagnostics ou thérapeutiques)



Discussion

- Construction modèles de détection
 - Injection de nombreuses données médicales
 - 2497 variables explicatives
 - Pertinences variables (Codes CIM-10, données microbiologiques (micro-organismes), données Biologiques (CRP), données administratives (durée hospitalisation > 5 jours)
 - Sélection a priori diminue performance modèle de détection
 - Meilleures performances
 - Sensibilités, spécificités, VPP, VPN supérieures
 - AUC proche de 1
 - NRI positifs

Discussion

- Codage données
 - ▣ Identification codes intéressants
 - Evènement infectieux potentiel après la chirurgie
 - T814, L022, QZJA011
 - ▣ Amélioration codage
 - ▣ Nombreuses études basées sur l'apport des données du PMSI
 - Infections associées aux soins
 - Diverses approches (1)
 - Text-mining (2)

(1) Brossette SE, Sprague AP, Hardin JM, Waites KB, Jones WT, Moser SA. Association rules and data mining in hospital infection control and public health surveillance. J Am Med Inform Assoc JAMIA. 1998 Aug;5(4):373-81.
(2) Campillo-Gimenez B, Garcelon N, Jarno P, Chaplain JM, Cuggia M. Full-text automated detection of surgical site infections secondary to neurosurgery in Rennes, France. Stud Health Technol Inform. 2013;192:572-5.

Merci pour votre attention