

Prédiction de l'incidence des infections du site opératoire (ISO) par un modèle hiérarchique

L. Saunders (Interne de santé publique, CCLIN Ouest)

M. Perennec-Olivier (Epidémiologiste, CCLIN Ouest)

Pr J-F. Viel (PU-PH, Département de Santé Publique)

Dr P. Jarno (PH, CCLIN Ouest)

La surveillance des ISO en France

- **Définition épidémiologique d'une ISO¹** : infection nosocomiale survenant dans les 30 jours suivant une intervention chirurgicale
- **Incidence globale² \approx 1%**
- **Surveillance nationale** coordonnée par le RAISIN depuis 2001 (1 507 690 patients de 2006 à 2011)
- Un indicateur de surveillance : le **Ratio Standardisé d' Incidence (RSI)³**

$$\text{RSI} = \frac{\text{nombre d'ISO observé}}{\text{nombre d'ISO attendu}}$$

$$\hat{a}_{i=1}^{\text{Nbre de patients}} \left[\frac{\exp(\text{Logit}P_i)}{1 + \exp(\text{Logit}P_i)} \right]$$

Modèle de régression logistique
(Données 2004-2008)

¹ RAISIN (Réseau d'Alerte, d'Investigation et de Surveillance des Infections Nosocomiales). Protocole national 2012. Surveillance des infections du site opératoire. 2011

² InVS. Surveillance des infections du site opératoire en France en 2009-2010. Résultats. 2012

³ Rioux C. The standardized incidence ratio as a reliable tool for surgical site infection surveillance. Infection control and hospital epidemiology. 2006;27(8):817-24.

Comment prédire le risque individuel d'ISO?

- A partir de **facteurs de risque individuels** (âge, sexe, durée d'intervention...)
- **Qu'en est-il des facteurs collectifs?**
 - Service, établissement hospitalier

Modèle de régression logistique multi-niveau

Régression logistique

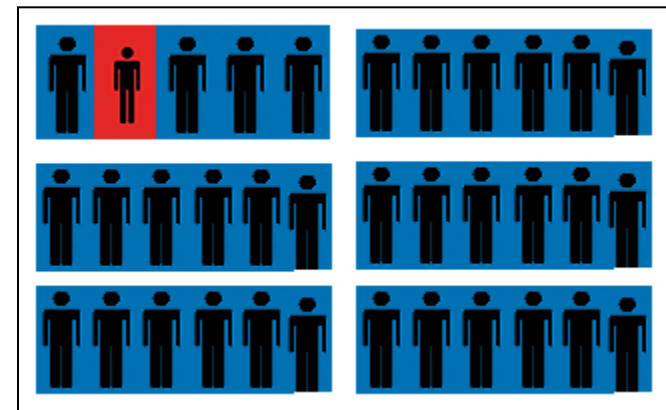
Patient



Service



Etablissement



Objectifs

- **Construire un modèle prédictif de survenue d'une ISO chez un patient hospitalisé**
- **Evaluer la performance prédictive du modèle sur un échantillon indépendant**

Population d'étude

■ Source

Base de données nationale de surveillance des ISO de l'année 2011

■ Inclusion

Tous les patients opérés dans les services volontaires participant à la surveillance (soit 330 015 patients)

■ Echantillon test

- Tirage aléatoire : 20% parmi les établissements de l'année 2011 (n = 756)



151 établissements, 502 services, 62 280 patients

Recueil de données

| | | |
|----------------------|--|--|
| PATIENT | Présence ou non d'une ISO dans les 30 jours post opératoires | |
| | Caractéristiques socio démographiques et cliniques | Age Sexe ASA |
| | Caractéristiques de l'intervention | Type Chirurgie carcinologique Classe de contamination Durée d'intervention Intervention en urgence ou programmée Intervention en ambulatoire ou non Durée de suivi post opératoire |
| | En cas d'ISO | Caractéristiques cliniques |
| SERVICE | | - |
| ETABLISSEMENT | Caractéristiques de l'établissement | Statut Type d'établissement |

Méthode d'analyse statistique

■ Modèle multi-niveau de régression logistique

- Variable expliquée : **ISO (présence/absence)**
- Variables explicatives réparties selon **3 niveaux**

Patient
Service
Etablissement

■ Etapes

- Modèle « vide » à intercept aléatoire
- Analyse univariée multi-niveau
- Analyse multivariée multi-niveau :
 - Modèle à intercept aléatoire avec caractéristiques individuelles (sélection des variables significatives au seuil $p < 0,20$ en analyse univariée),
 - Modèle à intercept aléatoire avec caractéristiques individuelles et pente(s) aléatoire(s).

Etapes de construction du modèle

■ Modèle « vide »

$$\begin{aligned}\text{logit}(\pi_{ijk}) &= \beta_{0jk} \\ \beta_{0jk} &= \beta_0 + v_{0k} + u_{0jk}\end{aligned}$$

Intercept aléatoire

Effets aléatoires :

- Etablissement (v_{0k})
- Service (u_{0jk})

■ Modèle avec facteurs individuels

$$\begin{aligned}\text{logit}(\pi_{ijk}) &= \beta_{0jk} + \beta_1 x1_{ijk} + \beta_2 x2_{ijk} \\ \beta_{0jk} &= \beta_0 + v_{0k} + u_{0jk}\end{aligned}$$

β_1 et β_2 : paramètres de régression fixes associés à x_1 et x_2

■ Modèle avec facteurs individuels et pente aléatoire

$$\begin{aligned}\text{logit}(\pi_{ijk}) &= \beta_{0jk} + \beta_{1j} x1_{ijk} + \beta_2 x2_{ijk} \\ \beta_{0jk} &= \beta_0 + v_{0k} + u_{0jk} \\ \beta_{1j} &= \beta_1 + u_{1jk}\end{aligned}$$

β_{1j} : paramètre de régression aléatoire associé à x_1

u_{1jk} : effet aléatoire lié à la variable x_1

Calcul de l'Odds Ratio Médian (MOR)

■ **Définition** : odds ratio médian entre le service à risque d'ISO le plus élevé et le service à risque d'ISO le moins élevé, lorsqu'on tire au sort plusieurs paires de services (méthode de Monte Carlo).

■ **Formule** ^{1,2}

$$MOR = \exp\left[\sqrt{2 * \sigma^2 u_0} * 0,6745\right]$$

■ **Interprétation**

- MOR = 1 → pas d'hétérogénéité entre services
- MOR > 1 → il existe une hétérogénéité entre les services

¹ Merlo J, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health. 2006;60(4):290-7.

² Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. Am. J. Epidemiol. 2005 1;161(1):81-8.

Méthode de validation

Performance prédictive du modèle?

Echantillon de validation externe
n = 51 348 (20% des établissements)

Coefficients issus du
modèle multi-niveau

Coefficients issus du
modèle de référence
(régression logistique sur les
données nationales 2004-2008)

Appliqués sur les données des patients
de l'échantillon de validation

Appliqués sur les données des patients
de l'échantillon de validation

Courbe ROC¹ et AUC²

Courbe ROC et AUC

**Différence significative
entre les 2 AUC?**

¹ ROC : Receiver Operating Characteristics

² AUC : Area Under Curve

ANALYSE DESCRIPTIVE

- **Incidence globale des ISO = 1%** (623 / 62 280)
 - 40% diagnostiquées avant la sortie
 - 48,1% superficielles
 - Délai moyen de survenue (\pm écart-type) = 11,9 (\pm 7,8) jours
 - Taux de décès = 1,6%
- **Sex-ratio = 0,76**
- **Age moyen (\pm écart-type) = 56,2 (\pm 21,5) ans**

Modèle « vide »

Variances des effets aléatoires par niveau (modèle « vide »)

| Niveau | σ^2 * | P |
|---------------|--------------|------------|
| Patient | 2,95 | $<10^{-9}$ |
| Service | 1,64 | $<10^{-9}$ |
| Etablissement | 0,19 | 0,09 |

* σ^2 = Variances des effets aléatoires liés à chaque niveau

MOR = 3,02 [IC 95% 2,47 ; 3,71]

-
- Seuls les 2 niveaux « patient » et « service » sont conservés dans le modèle
 - Sans prendre en compte les facteurs individuels, le risque individuel médian est multiplié par 3,02 entre le service à risque le moins élevé et le service à risque le plus élevé.

Facteurs d' ISO en analyse multivariée multi-niveau

| Caractéristiques patient | p |
|---|-------------|
| Hommes | 0,04 |
| Age \geq 65 ans | 0,19 |
| Score ASA \geq 3 | $< 10^{-8}$ |
| Durée de suivi $<$ 15 jours | $< 10^{-9}$ |
| Durée de séjour préopératoire \geq 48 heures | $< 10^{-3}$ |
| Chirurgie « contaminée » ou « sale » | $< 10^{-6}$ |
| Absence de coeliochirurgie | $< 10^{-3}$ |
| Chirurgie digestive | $< 10^{-9}$ |
| Durée d'intervention $>$ 75 ^{ème} percentile | $< 10^{-9}$ |
| Intervention en urgence | 0,52 |
| Intervention en milieu hospitalier | $< 10^{-8}$ |

Modèle à 2 niveaux, avec caractéristiques individuelles

Variances des effets aléatoires par niveau (modèle multivarié avec caractéristiques individuelles)

| Niveau | σ^2 * | P |
|---------|--------------|------------|
| Patient | 3,65 | $<10^{-9}$ |
| Service | 1,82 | $<10^{-9}$ |

* σ^2 = Variances des effets aléatoires liés à chaque niveau

$$\text{MOR} = 3,55 \text{ [IC 95\% 2,99 ; 4,29]}$$

- En tenant compte des caractéristiques individuelles, le risque individuel médian est multiplié par 3,55 entre le service à risque le moins élevé et le service à risque le plus élevé

Modèle à 2 niveaux, avec caractéristiques individuelles et pente aléatoire sur la variable « durée de suivi »

Variances des effets aléatoires par niveau (modèle final avec pente aléatoire)

| Niveau | σ^2 * | P |
|---------|--------------|------------|
| Patient | 2,91 | $<10^{-9}$ |
| Service | 3,50 | $<10^{-9}$ |

* σ^2 = Variances des effets aléatoires liés à chaque niveau

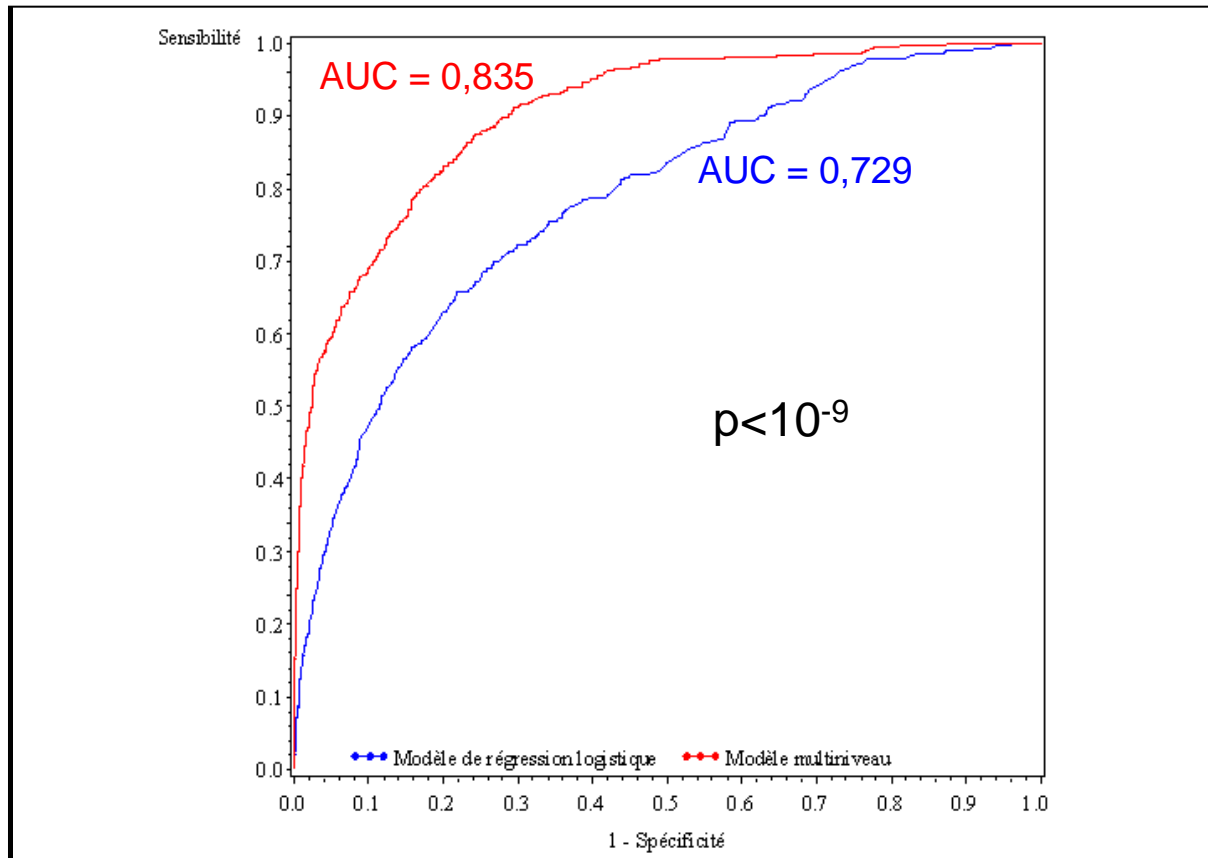
Suivi < 15 jours : MOR = 7,53 [IC 95% 5,42 ; 10,46]
Suivi \geq 15 jours : MOR = 3,21 [IC 95% 2,32 ; 4,45]

→ **L' hétérogénéité du risque ISO entre services est plus marquée si le suivi est < 15 jours**

Interprétation :

- La variable « durée de suivi » individuelle reflète probablement une politique de service en termes de suivi post hospitalisation
- Alors cette hétérogénéité plus importante dans les 15 premiers jours reflète probablement une hétérogénéité entre services ayant une procédure de suivi systématique et services sans procédure de suivi

COMPARAISON DES COURBES ROC



Le modèle multi-niveau possède une performance prédictive nettement supérieure au modèle de régression logistique de référence

Discussion : données et méthode statistique

- Des données de bonne qualité (< 1% de données manquantes, cohérence)
- Des variables individuelles pertinentes, reconnues dans la littérature scientifique et utilisées dans les systèmes de surveillance américains et européens^{1,2,3}
- Une puissance statistique suffisante⁴
- Limite : variable service non exploitée (procédure de suivi)
- **Modèles multi-niveau**
 - **Avantages**
 - Meilleure estimation et meilleure précision des paramètres
 - Prise en compte de la non indépendance entre patients d' un même service
 - Identification d' un effet contextuel : le service
 - **Limites**
 - Complexité de mise en œuvre (choix des algorithmes, convergence)
 - Interprétation

¹ Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol.* 1992 oct;13(10):606-8.

² ECDC. Surveillance of surgical site infections in Europe, 2004-2008 [Internet]. 2012.

³ Astagneau P, L'Héritier F, Daniel F, Parneix P, Venier A-G, Malavaud S, et al. Reducing surgical site infection incidence through a network: results from the French ISO-RAISIN surveillance system. *J. Hosp. Infect.* 2009 ;72(2):127-34.

⁴ Moineddin R, Matheson FI, Glazier RH. A simulation study of sample size for multilevel logistic regression models. *BMC Med Res Methodol.* 16 2007;7:34

Perspectives

- Un apport important en termes d'**explicitation du risque ISO**
 - Mise en évidence de l'importance de l'effet service
 - Meilleure estimation du nombre d'ISO théorique pour le calcul du RSI (cliniciens)
- **Le modèle multi-niveau devrait à terme se substituer au modèle de régression logistique actuel**
- Nécessité d'étudier plus précisément l'**effet du service**
 - Réaliser une étude complémentaire rétrospective (recueil d'informations concernant l'organisation du suivi après l'hospitalisation)
 - Renseigner systématiquement la variable « procédure de suivi »
 - Au-delà du suivi post-hospitalisation : compléter le recueil par des données concernant les services (organisation, pratiques, etc.)

Merci de votre attention