Big data et EI : intérêt et limites

Pr Xavier Duval

Hôpital Bichat Claude Bernard, Paris

JNI 24es Journées Nationales d'Infectiologie

Grenoble
et la région Auvergne-Rhône-Alpes
ALPEXPO
du mercredi 7 au vendredi 9 juin 2023

SPILF — SOCIÉTÉ DE PATHOLOGIE INFECTIEUSE DE LANGUE FRANÇAISE

CMiT — Collège des universitaires de Maladies Infectieuses et Tropicales

AEPEI

lame UMR 1137

Instituts thématiques
Inserm
Institut national
de la santé et de la recherche médicale

JNI
24es Journées Nationales d'Infectiologie

Grenoble
et la région Auvergne-Rhône-Alpes
ALPEXPO
du mercredi 7 au vendredi 9 juin 2023

Déclaration de liens d'intérêt avec les industries de santé en rapport avec le thème de la présentation (loi du 04/03/2002) :

Intervenant : Duval Xavier

Titre : Big data et EI : intérêt et limites

L'orateur ne souhaite pas répondre ☐

Consultant ou membre d'un conseil scientifique — OUI ☐ NON ☑

Conférencier ou auteur/rédacteur rémunéré d'articles ou documents — OUI ☐ NON ☑

Prise en charge de frais de voyage, d'hébergement ou d'inscription à des congrès ou autres manifestations — OUI ☐ NON ☑

Investigateur principal d'une recherche ou d'une étude clinique — OUI ☐ NON ☑

Endocardite infectieuse

- **Maladie rare**

- **Absence de surveillance épidémiologique**
  - Maladie non à déclaration obligatoire
  - Non suivie dans EPIBAC (contrairement aux méningites bactériennes)

- **Pas de structuration du soin en CNR**

- **Recherches observation[nelle] ou intervention[nelle] compliquées**
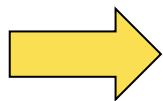
  ⟹ **Intérêt des bases médico administratives ?**

Bases de données médico administratives

- **Bases de données médico administratives**
  - Constituées à des fins économiques
    - Remboursement des couts de santé dont établissements de santé
    - Estimation de la performance des médecins par les caisses privées
      - Adresser leurs clients aux médecins avec meilleurs rapports couts/bénéfices
  - Informations en partie déclaratives

Bases de données médico administratives France

- **Entrepôts de données de santé des hôpitaux**
  - Informations uniquement de l'établissement
  - Richesse de la base dépend des informations « qui remontent »
    - Observation médicale
    - Compte rendu d'hospitalisation
    - Données biologiques
    - Compte rendu d'imagerie
    - Codage PMSI (Résumé versus détaillé …)
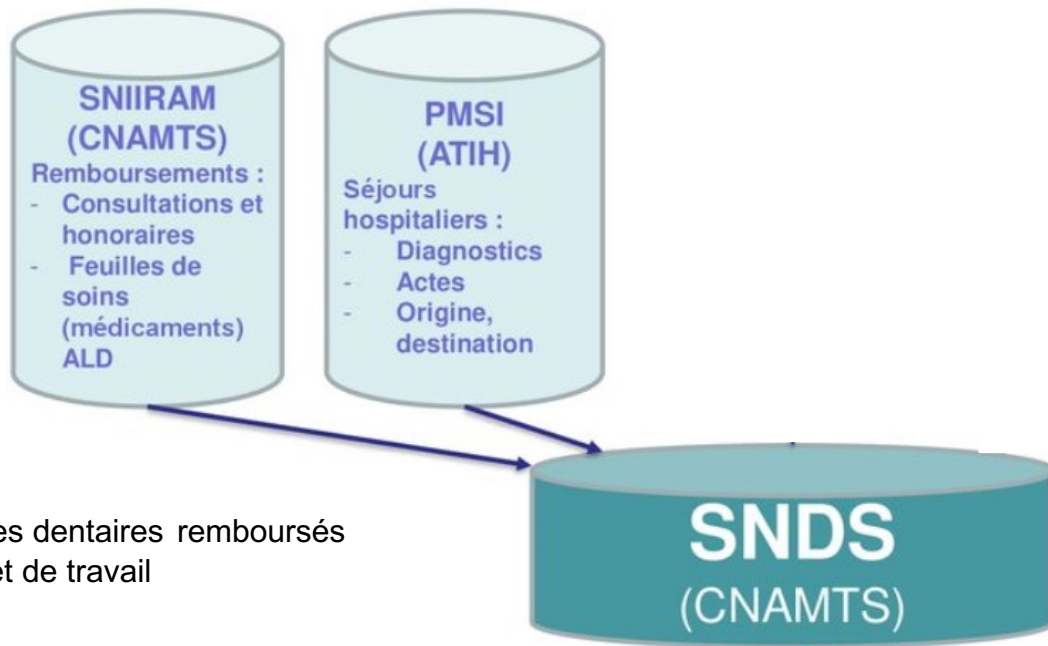 - Analyse textuelle complexe (double négation)

Plus ou moins structurées

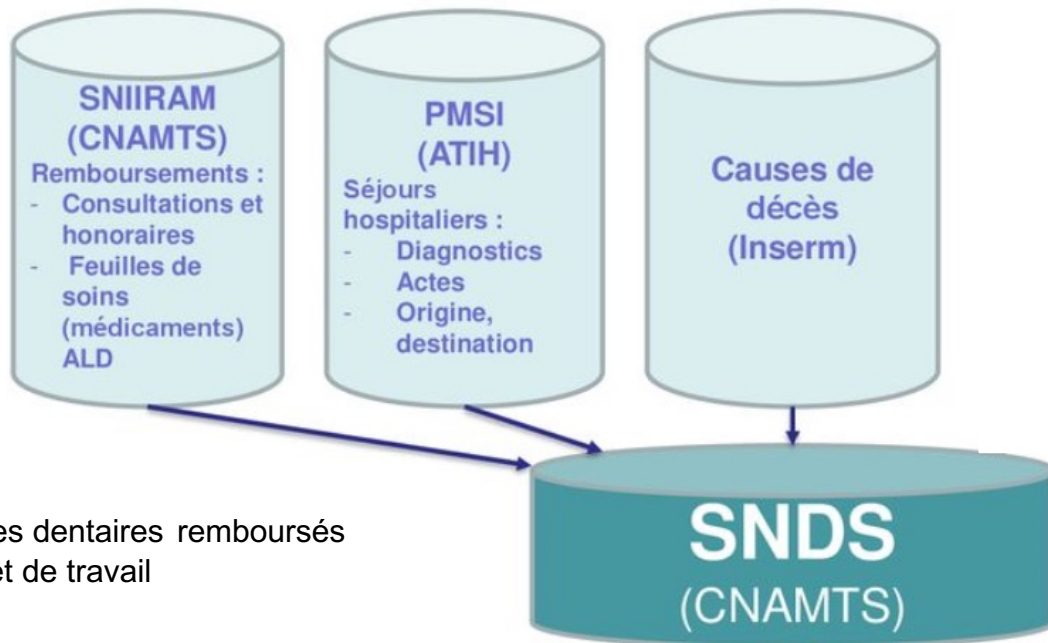# Architecture Système National des Données de Santé SNDS

SNIIRAM (CNAMTS)
Remboursements :
- Consultations et honoraires
- Feuilles de soins (médicaments) ALD

Actes dentaires remboursés
Arrêt de travail

SNDS (CNAMTS)

Source : Philippe Tuppin, Direction de la Stratégie, des Études et des Statistiques, Données de santé : du SNIIRAM au SNDS
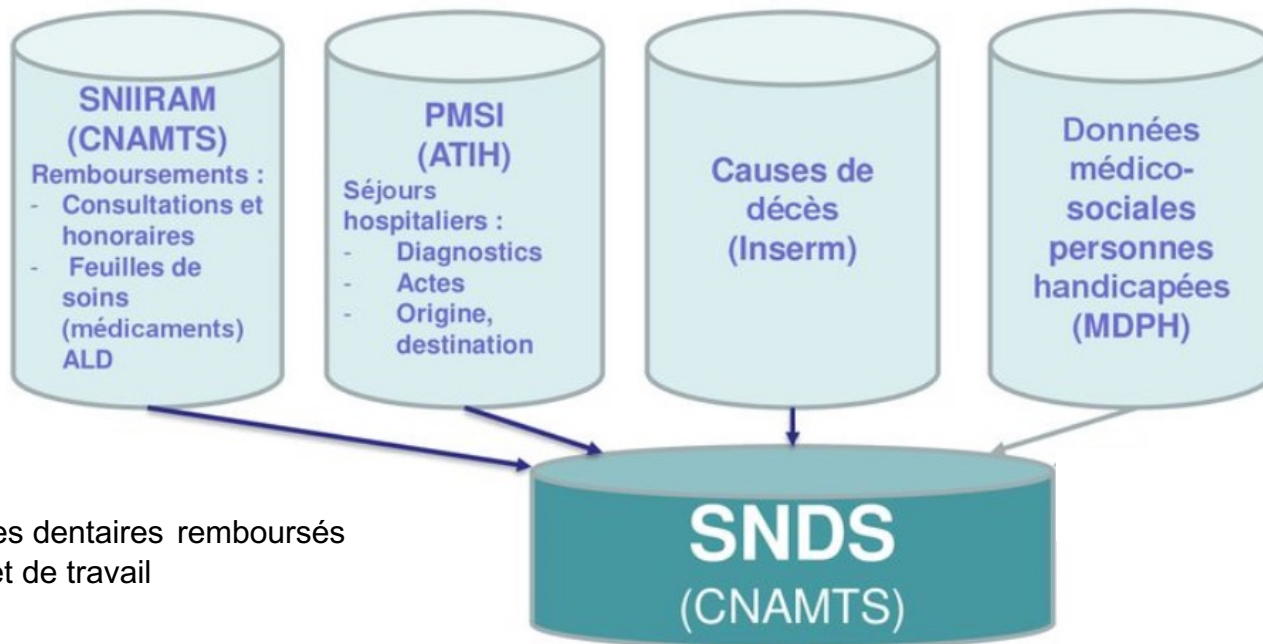
Architecture Système National des Données de Santé SNDS

Actes dentaires remboursés
Arrêt de travail

Architecture Système National des Données de Santé SNDS

Actes dentaires remboursés
Arrêt de travail

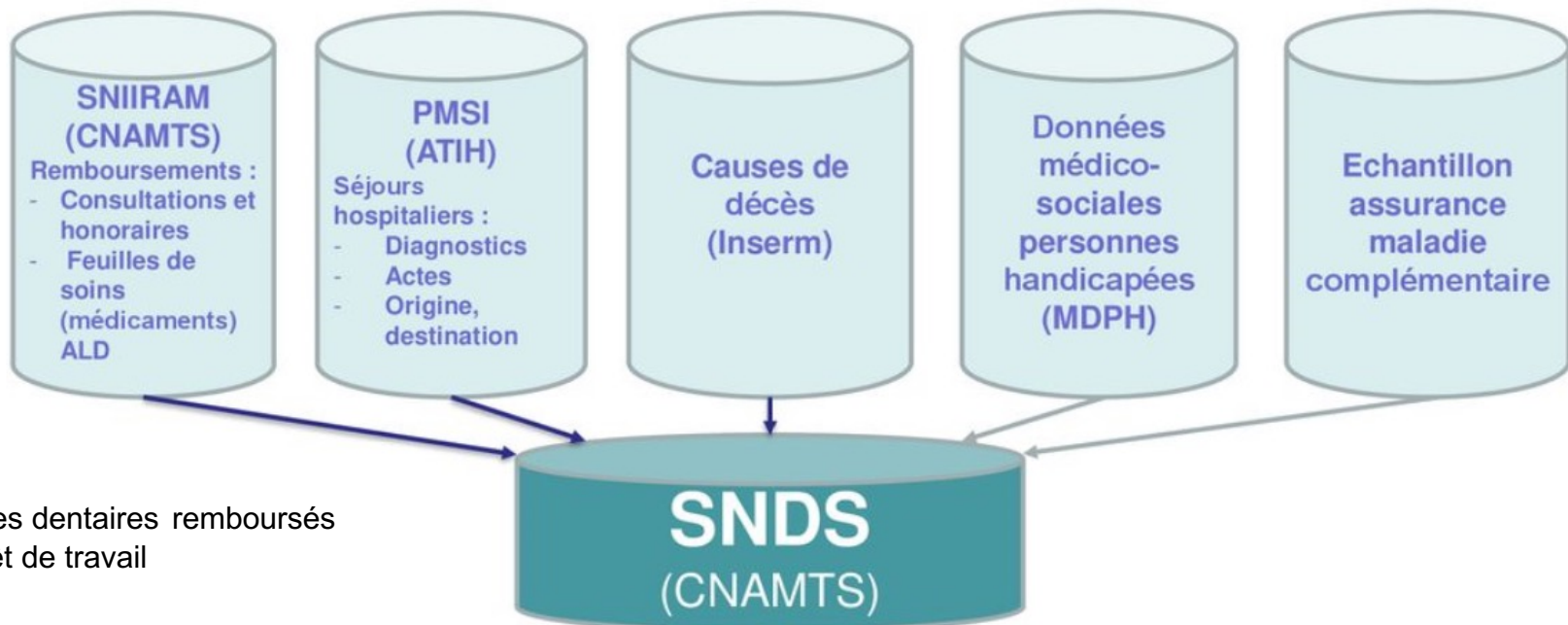# Architecture Système National des Données de Santé SNDS

Actes dentaires remboursés
Arrêt de travail

Architecture Système National des Données de Santé SNDS

Actes dentaires remboursés
Arrêt de travail

Source : Philippe Tuppin, Direction de la Stratégie, des Études et des Statistiques, Données de santé : du SNIIRAM au SNDS

Identification des caractéristiques des PTS

- **PMSI: DP, DS, DA**
  - ATCD pose ou remplacement prothèses valvulaires cardiaques
  - PM, Défibrillateur
- **Algorithme comorbidités**

  Codes d'hospitalisation

  Remboursement Traitement: ADO, Tt HTA, …..

  100%

Bases de données médico administratives France

- **Indice de précarité**

- **Département de domicile**

- **Absence de données sur traitement hospitalier**

- **Réglementairement**

  - Accord CNIL et comité d'éthique

  - Possibilité de chainage à des données semi-nominatives de recherche mais complexe, nécessitant tiers de confiance (Consentement écrit du patient)

Bases de données médico administratives France

- **Identification Endocardite par codage ICD 10 / 9**

**Table 1** ICD-10 Endocarditis codes and corresponding ICD-9 codes (and clinical modifications)

ICD-10 Code	Description	Corresponding ICD-9 Code/ICD-9-CM code	Description
Included			
I33 (I330)	Acute and subacute infective endocarditis	4210	Acute and subacute infective endocarditis
I38	Endocarditis, valve unspecified	4249	Endocarditis valve unspecified cause
(I38.X)		42499	Other endocarditis valve unspecified
I339	Acute and subacute endocarditis, unspecified	4219	Acute endocarditis unspecified
T826	Infection and inflammatory reaction due to cardiac valve prosthesis	99661	Infection and inflammatory reaction due to cardiac device implant and graft
B376	Candidal endocarditis	11281	Candidal endocarditis
I39	Endocarditis and heart valve disorders in diseases classified elsewhere	11504	*Histoplasma capsulatum* endocarditis
(I390)			Histoplasma duboisii endocarditis
		11514	Histoplasmosis endocarditis
		11594	Acute and subacute infective endocarditis in diseases classified elsewhere
		4211	Endocarditis in diseases classified elsewhere
I398	Endocarditis, valve unspecified, in diseases classified elsewhere	42491	

Description épidémiologique

BDMA et Description épidémiologique EI

- **Incidence et évolution au cours du temps (**
  - Age, sexe …..
- **Comorbidités**
  - Prédispositions cardiaques
  - Diabète …
- **Caractéristiques de l'EI**
  - Microorganismes
  - Pronostic …..

Figure 1. Standardized Incidence of Infective Endocarditis in California and New York State From 1998 Through 2013
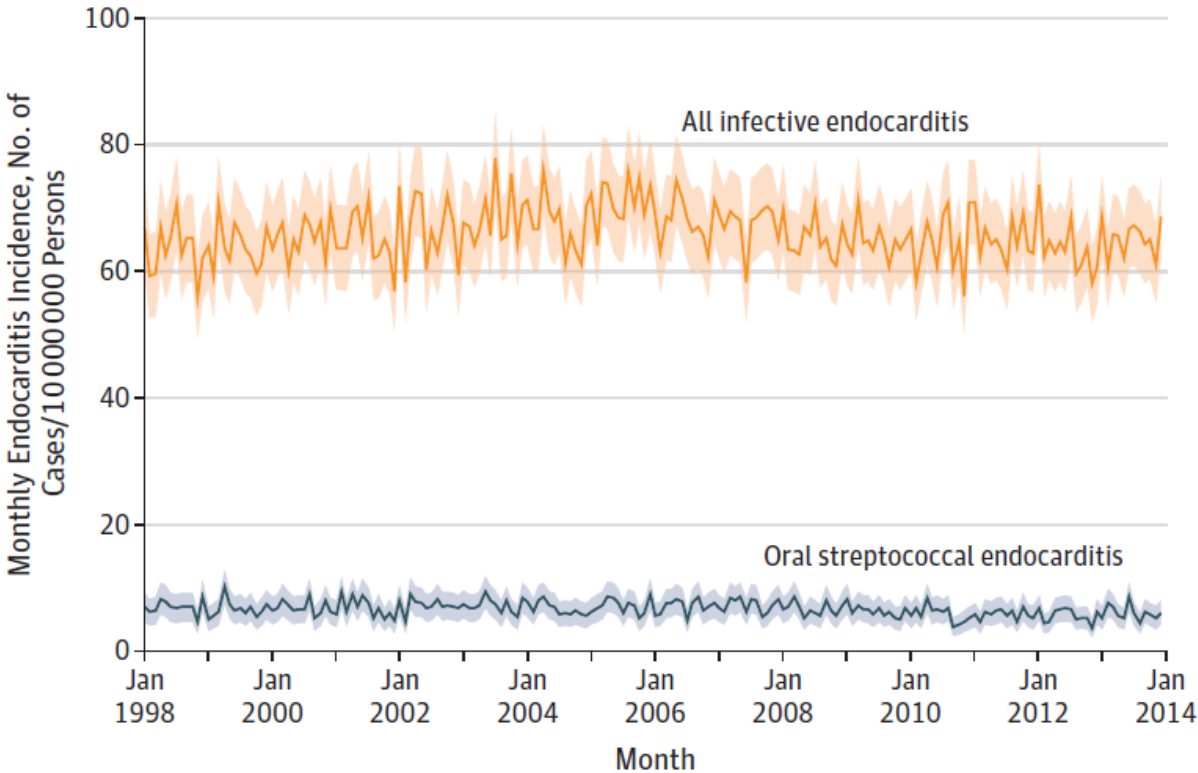
BDMA et Description épidémiologique

Incidence of infective endocarditis in patients considered at moderate risk

Lauge Østergaard[1]*, Nana Valeur[2], Andrew Wang[3], Henning Bundgaard[1], Mohsin Aslam[1], Gunnar Gislason[4,5], Christian Torp-Pedersen[6], Niels Eske Bruun[4,6,7], Lars Søndergaard[1], Lars Køber[1], and Emil Loldrup Fosbøl[1]

Incidence of infective endocarditis among patients considered at high risk

Lauge Østergaard[1]*, Nana Valeur[2], Nikolaj Ihlemann[1], Henning Bundgaard[1], Gunnar Gislason[3,4], Christian Torp-Pedersen[5,6], Niels Eske Bruun[7,8], Lars Søndergaard[1], Lars Køber[1], and Emil Loldrup Fosbøl[1]

Quantifying infective endocarditis risk in patients with predisposing cardiac conditions

Martin H. Thornhill[1,2]*, Simon Jones[3,4], Bernard Prendergast[5], Larry M. Baddour[6], John B. Chambers[5], Peter B. Lockhart[2], and Mark J. Dayer[7]

Incidence annuelle EI

/ 100 000	Ostergaard L (Eur H J 2018) Ostergaard L (Eur H J 2019)	Thornhill (Eur H J 2018)
ATCD EI	1 610	1 436
Valve prothétique	670	453
Cardiopathie congénitale	79	
PM - Défibrillateurs	189	68
Valvulopathie	156	280
Cardiomyopathies hypertrophiques	71	167
Population générale	16	3.62

Ostergaard L , Eur Heart J 2018
Thornhill MH, Eur Heart J. 2017
Ostergaard L , Eur Heart J 2019

Trends in Infective Endocarditis in California and New York State, 1998-2013

Nana Toyoda, MD; Joanna Chikwe, MD; Shinobu Itagaki, MD, MS; Annetine C. Gelijns, PhD;
David H. Adams, MD; Natalia N. Egorova, PhD

Table. Patient Characteristics in California and New York State From 1998 Through 2013, Overall and Trends by Years

Characteristic	No. (%) Overall (N = 75 829)
Demographic	
Age, mean (SD), y	62.3 (18.9)
Male	44 804 (59.1)
Race	
White	52 776 (69.6)
Black	10 074 (13.3)
Other	12 979 (17.1)
Comorbidities	
Hypertension	47 604 (62.8)
Complicated diabetes	11 165 (14.7)
Coronary artery disease	30 326 (40.0)
Peripheral vascular disease	11 763 (15.5)
Chronic obstructive pulmonary disease	19 185 (25.3)
Chronic kidney disease	22 308 (29.4)
Dialysis dependent	13 989 (18.4)
Liver disease	15 477 (20.4)
Cancer	13 121 (17.3)
Human immunodeficiency virus[c]	715 (1.7)
Predisposing factor	
History of congenital heart disease	3437 (4.5)
History of valve surgery	10 591 (14.0)
History of implanted pacemaker or defibrillator	9765 (12.9)

Toyoda N, JAMA 2017

Comorbiditées

Table. Patient Characteristics in California and New York State From 1998 Through 2013, Overall and Trends by Years

Characteristic	No. (%) Overall (N = 75 829)	1998-2001 (n = 16 511)	2002-2005 (n = 18 887)	2006-2009 (n = 19 611)	2010-2013 (n = 20 820)	Annual Percentage Change, % (95% CI)	P Value[a]
Demographic							
Age, mean (SD), y	62.3 (18.9)	61.0 (19.4)	61.9 (19.0)	63.0 (18.5)	63.3 (18.5)		<.001[b]
Male	44 804 (59.1)	9550 (57.8)	10 971 (58.1)	11 672 (59.5)	12 611 (60.6)	0.5 (0.2 to 0.7)	<.001
Race							
White	52 776 (69.6)	11 750 (71.2)	13 262 (70.2)	13 534 (69.0)	14 230 (68.3)	−0.3 (−0.5 to −0.2)	<.001
Black	10 074 (13.3)	2189 (13.3)	2547 (13.5)	2714 (13.8)	2624 (12.6)	−0.4 (−0.8 to 0.03)	.07
Other	12 979 (17.1)	2572 (15.6)	3078 (16.3)	3363 (17.1)	3966 (19.0)	1.7 (1.3 to 2.1)	<.001
Comorbidities							
Hypertension	47 604 (62.8)	8176 (49.5)	11 495 (60.9)	13 189 (67.3)	14 744 (70.8)	2.8 (2.6 to 3.0)	<.001
Complicated diabetes	11 165 (14.7)	1921 (11.6)	2632 (13.9)	3097 (15.8)	3515 (16.9)	3.0 (2.6 to 3.4)	<.001
Coronary artery disease	30 326 (40.0)	5589 (33.9)	7433 (39.4)	8303 (42.3)	9001 (43.2)	2.0 (1.7 to 2.2)	<.001
Peripheral vascular disease	11 763 (15.5)	2246 (13.6)	3295 (17.4)	3145 (16.0)	3077 (14.8)	0.3 (−0.1 to 0.7)	.11
Chronic obstructive pulmonary disease	19 185 (25.3)	3327 (20.2)	4675 (24.8)	5450 (27.8)	5733 (27.5)	2.5 (2.2 to 2.8)	<.001
Chronic kidney disease	22 308 (29.4)	3059 (18.5)	4397 (23.3)	6776 (34.6)	8076 (38.8)	6.3 (6.0 to 6.6)	<.001
Dialysis dependent	13 989 (18.4)	2459 (14.9)	3797 (20.1)	4000 (20.4)	3733 (17.9)	1.0 (0.6 to 1.3)	<.001
Liver disease	15 477 (20.4)	2544 (15.4)	3625 (19.2)	4108 (20.9)	5200 (25.0)	3.9 (3.5 to 4.3)	<.001
Cancer	13 121 (17.3)	2480 (15.0)	3157 (16.7)	3538 (18.0)	3946 (19.0)	1.9 (1.5 to 2.3)	<.001
Human immunodeficiency virus[c]	715 (1.7)	182 (2.0)	203 (1.9)	178 (1.6)	152 (1.3)	−3.4 (−4.9 to −1.8)	<.001
Predisposing factor							
History of congenital heart disease	3437 (4.5)	658 (4.0)	857 (4.5)	1043 (5.3)	879 (4.2)	0.7 (−0.01 to 1.5)	.05
History of valve surgery	10 591 (14.0)	2116 (12.8)	2515 (13.3)	2800 (14.3)	3160 (15.2)	1.6 (1.2 to 2.0)	<.001
History of implanted pacemaker or defibrillator	9765 (12.9)	1446 (8.8)	2207 (11.7)	2865 (14.6)	3247 (15.6)	4.8 (4.3 to 5.2)	<.001

Toyoda N, JAMA 2017

BDMA et Description épidémiologique

Trends in Infective Endocarditis in California and New York State, 1998-2013

Nana Toyoda, MD; Joanna Chikwe, MD; Shinobu Itagaki, MD, MS; Annetine C. Gelijns, PhD; David H. Adams, MD; Natalia N. Egorova, PhD

Characteristic	No. (%) Overall (N = 75 829)	1998-2001 (n = 16 511)	2002-2005 (n = 18 887)	2006-2009 (n = 19 611)	2010-2013 (n = 20 820)	Annual Percentage Change, % (95% CI)	P Value[a]
Disease type							
Native-valve endocarditis	54 332 (71.7)	12 299 (74.5)	13 747 (72.8)	14 038 (71.6)	14 248 (68.4)	−0.7 (−0.9 to −0.5)	<.001
Prosthetic-valve endocarditis	9777 (12.9)	1989 (12.0)	2355 (12.5)	2558 (13.0)	2875 (13.8)	1.3 (0.8 to 1.7)	<.001
Cardiac device–related endocarditis	2236 (3.0)	217 (1.3)	454 (2.4)	717 (3.7)	848 (4.1)	8.8 (7.8 to 9.9)	<.001
Drug abuse–related endocarditis	9484 (12.5)	2006 (12.1)	2331 (12.3)	2298 (11.7)	2849 (13.7)	0.9 (0.4 to 1.3)	<.001
Mode of acquisition							
Community-acquired	35 701 (47.1)	8288 (50.2)	8571 (45.4)	8683 (44.3)	10 159 (48.8)	−0.2 (−0.4 to 0.03)	.09
Health care–associated	40 128 (52.9)	8223 (49.8)	10 316 (54.6)	10 928 (55.7)	10 661 (51.2)	0.2 (−0.04 to 0.4)	.10
Nosocomial	13 304 (17.5)	2923 (17.7)	3475 (18.7)	3721 (19.0)	3185 (15.3)	−1.0 (−1.4 to −0.7)	<.001
Nonnosocomial	26 824 (35.4)	5300 (32.1)	6841 (36.2)	7207 (36.7)	7476 (35.9)	0.8 (0.5 to 1.1)	<.001

Micro organismes

| Characteristic	No. (%) Overall (N = 75 829)	1998-2001 (n = 16 511)	2002-2005 (n = 18 887)	2006-2009 (n = 19 611)	2010-2013 (n = 20 820)	Annual Percentage Change, % (95% CI)	P Value[a]
Causative organism							
Staphylococcus	29 172 (38.5)	6011 (36.4)	7424 (39.3)	7520 (38.3)	8217 (39.5)	0.6 (0.3 to 0.8)	<.001
Staphylococcus aureus	24 179 (31.9)	4786 (29.0)	6170 (32.7)	6272 (32.0)	6951 (33.4)	1.0 (0.7 to 1.3)	<.001
Methicillin-resistant	9161 (12.1)	878 (5.3)	2399 (12.7)	2886 (14.7)	2998 (14.4)	6.3 (5.8 to 6.8)	<.001
Methicillin-sensitive	15 018 (19.8)	3908 (23.7)	3771 (20.0)	3386 (17.3)	3953 (19.0)	−2.0 (−2.3 to −1.6)	<.001
Streptococcus	20 157 (26.6)	4321 (26.2)	4761 (25.2)	5257 (26.8)	5818 (27.9)	0.7 (0.4 to 1.0)	<.001
Oral streptococci	7640 (10.1)	1759 (10.7)	1970 (10.4)	2019 (10.3)	1892 (9.1)	−1.2 (−1.7 to −0.8)	<.001
Gram-negative bacilli	4235 (5.6)	897 (5.4)	1022 (5.4)	1186 (6.0)	1130 (5.4)	0.1 (−0.5 to 0.8)	.68
Fungus	1316 (1.7)	285 (1.7)	356 (1.9)	365 (1.9)	310 (1.5)	−1.0 (−2.2 to 0.2)	.10
Other	1745 (2.3)	386 (2.3)	389 (2.1)	456 (2.3)	514 (2.5)	0.8 (−0.2 to 1.9)	.12
Unspecified	19 204 (25.3)	4611 (27.9)	4935 (26.1)	4827 (24.6)	4831 (23.2)	−1.5 (−1.9 to −1.2)	<.001

Toyoda N, JAMA 2017

JAMA | Original Investigation

Trends in Infective Endocarditis in California and New York State, 1998-2013

Nana Toyoda, MD; Joanna Chikwe, MD; Shinobu Itagaki, MD, MS; Annetine C. Gelijns, PhD; David H. Adams, MD; Natalia N. Egorova, PhD

Figure 1. Standardized Incidence of Infective Endocarditis in California and New York State From 1998 Through 2013
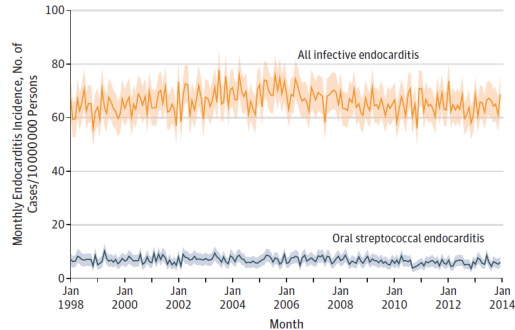
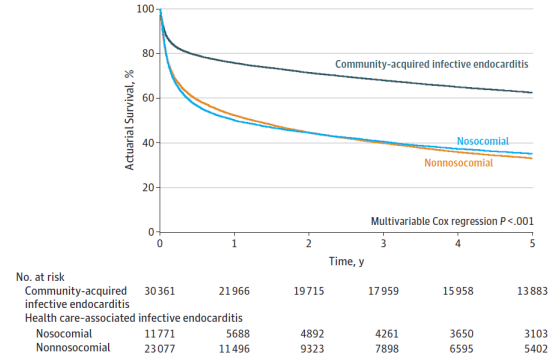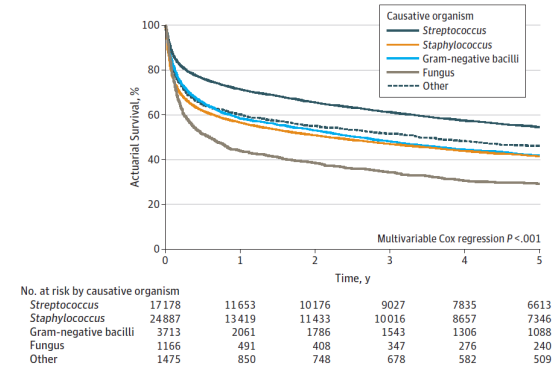Figure 2. Five-Year Survival of Patients With Infective Endocarditis, Stratified by Mode of Acquisition, in California and New York State

Figure 3. Five-Year Survival of Patients With Infective Endocarditis, Stratified by Pathogen, in California and New York State

Toyoda N, JAMA 2017

Recherche d'associations

BDMA et associations

- **Analyse descriptive**
  - Modification Incidence / Caractéristiques en fonction d'interventions
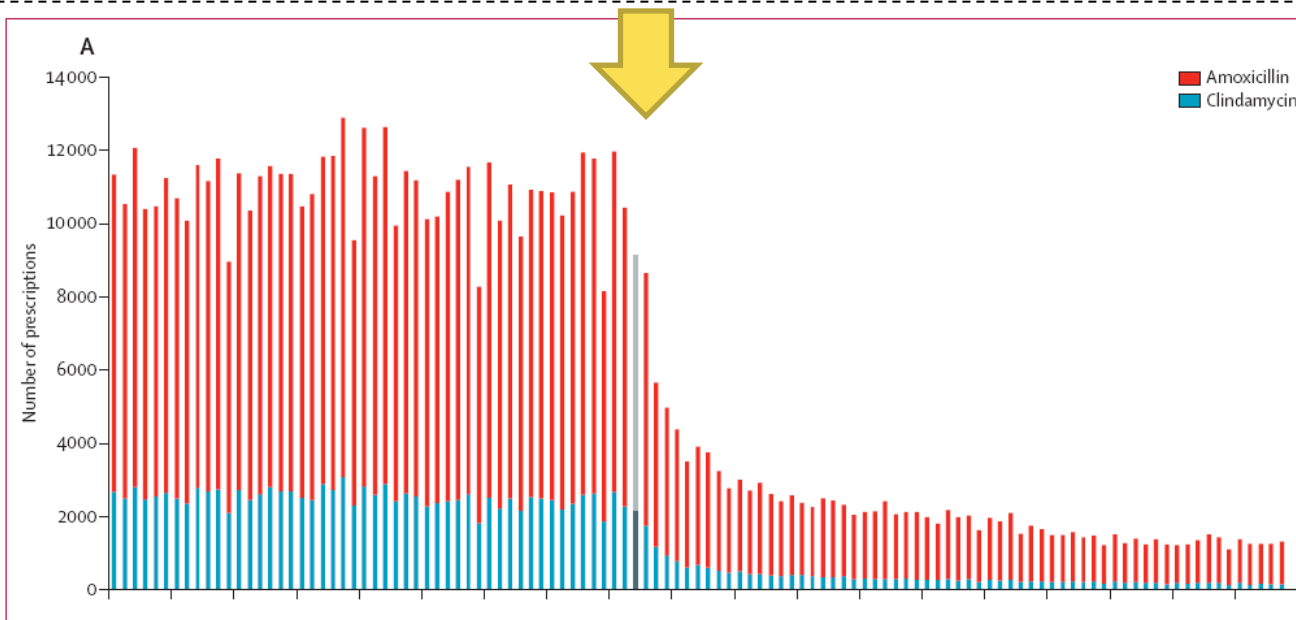   - Fait en particulier dans le cadre des modifications de prophylaxie de l'EI

BDMA et associations

- **Analyse des codes ICD**

Thornhill BMJ. 2011;342:d2392 - Fedeli BMC Infect Dis 2011, 11:48
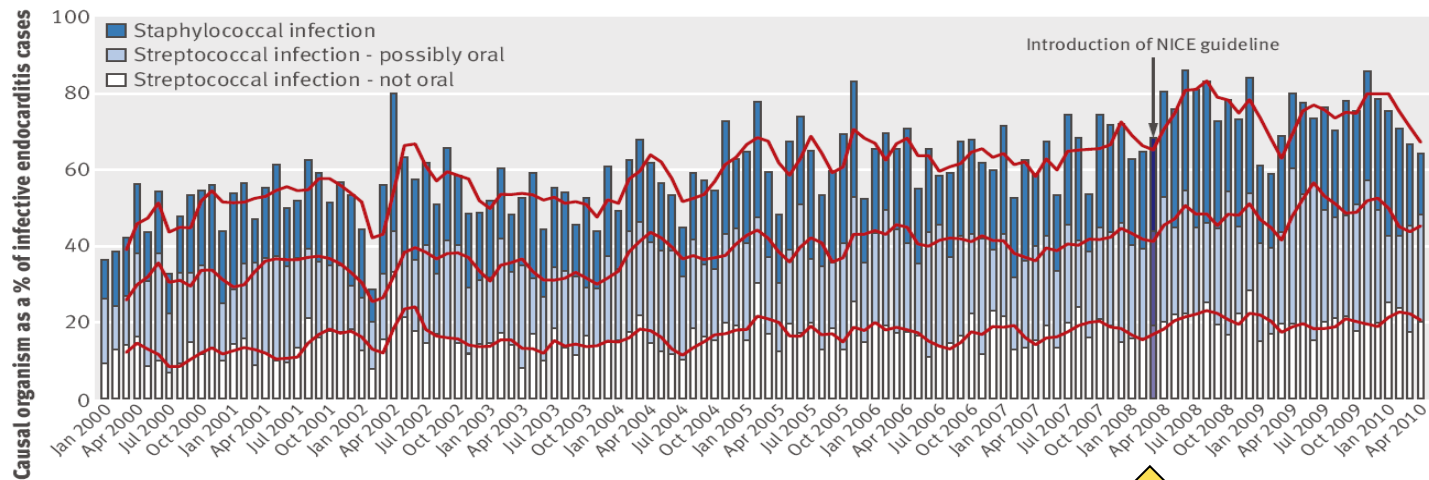
BDMA et associations

2008 NICE Guidelines: **abandon de la** prophylaxie antibiotique en Angleterre

90% réduction de la prophylaxie antibiotique après 2008

M Dayer Lancet 2015

24es JNI, GRENOBLE

Prevention

Figure 3: Incidence of infective endocarditis, by risk group

The figure shows the number of cases of infective endocarditis (superspells) recorded each month in individuals at high risk of developing infective endocarditis (solid red line) and those at lower risk (solid blue line). Data are corrected for change in the size of the total English population (not for change in the size of the high-risk or lower-risk groups). The vertical dashed line indicates March, 2008, the month in which cessation of antibiotic prophylaxis for infective endocarditis was recommended by the National Institute for Health and Clinical Excellence (NICE).[13] The trend lines for high-risk (dashed red line) and lower-risk (dashed blue line) individuals before and after introduction of the NICE guidelines are also shown.

Augmentation du nombre d'EI de 35 cas supplémentaires par mois

BDMA et associations

- **Analyse descriptive**
  - Modification Incidence / Caractéristiques en fonction d'interventions
    - Fait en particulier dans le cadre des modifications de prophylaxie de l'EI

- **Analyses causales (Case control study)**
  - Sarah Tubiana
  - Martin Thornhill
  - Imre Janszky

Etude de type Case-Crossover

- **Définition des périodes**

| Control period | Washout period | Control period | Washout period | Control period | Washout period | Case period | Oral streptococcal infective endocarditis |

Time (months preceding oral streptococcal infective endocarditis)

21   18   15   12   9   6   3   0

Période à risque

ORIGINAL INVESTIGATIONS
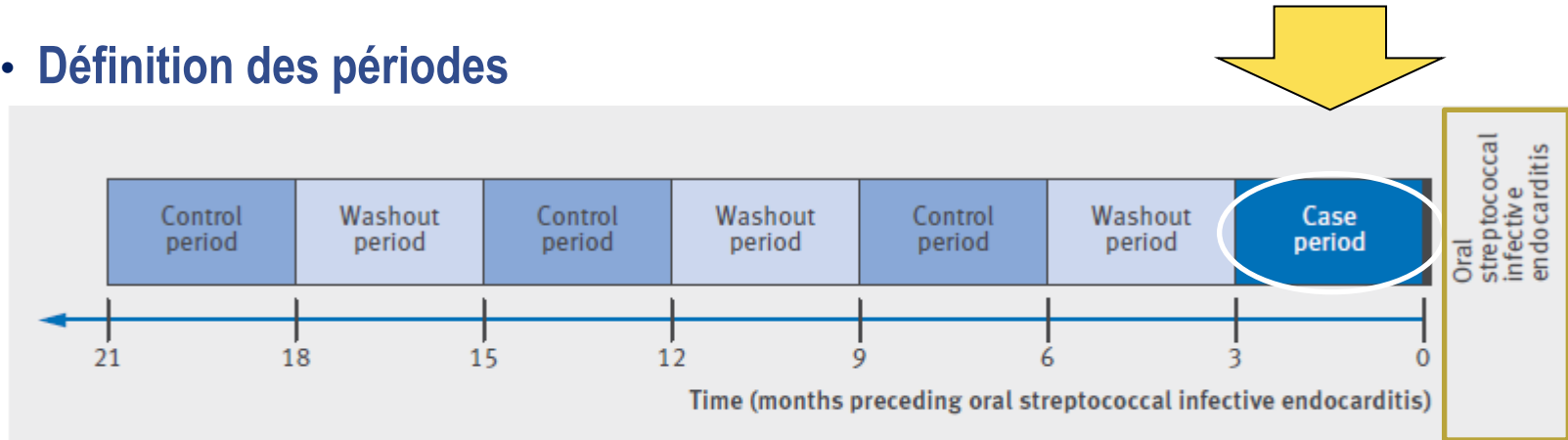
Invasive Procedures Associated With the Development of Infective Endocarditis

Imre Janszky, MD, PhD,[a,b] Katalin Gémes, PhD,[c] Staffan Ahnve, MD, PhD,[c] Hilmir Asgeirsson, MD, PhD,[d,e] Jette Möller, PhD[c]

Antibiotic Prophylaxis Against Infective Endocarditis Before Invasive Dental Procedures

Martin H. Thornhill, MBBS, BDS, PhD,[a,b] Teresa B. Gibson, PhD,[c] Frank Yoon, PhD,[c] Mark J. Dayer, MBBS, PhD,[d] Bernard D. Prendergast, BM, BS, DM,[e] Peter B. Lockhart, DDS,[b] Patrick T. O'Gara, MD,[f] Larry M. Baddour, MD[g]

RESEARCH

Dental procedures, antibiotic prophylaxis, and endocarditis among people with prosthetic heart valves: nationwide population based cohort and a case crossover study

Sarah Tubiana,[1,2] Pierre-Olivier Blotière,[2] Bruno Hoen,[3] Philippe Lesclous,[4] Sarah Millot,[5] Jérémie Rudant,[2] Alain Weill,[2] Joel Coste,[2] François Alla,[2] Xavier Duval[1]

24es JNI, GRENOBLE

24es JNI 2023 Grenoble

Etude de type Case-Crossover

- Analyse Case-crossover analysis (et étude de cohorte)

- 7 951 972 sujets américains du nord

- Sujets à haut risque d'EI

  - Procédures invasives dentaires :   OR: 2.00   95% CI: 1.59-2.52

  - Extraction dentaire: OR: 11.08   95% CI: 7.34-16.74

  - Antibioprophylaxie: OR: 0.49   95% CI: 0.29-0.85

Thornhill M, JAAC 2022

Etude de type Case-Crossover

Temporal association between invasive procedures and infective endocarditis

Martin H Thornhill [iD] ,[1,2] Annabel Crum,[3] Richard Campbell,[3] Tony Stone,[3] Ellen C Lee,[3] Mike Bradburn,[4] Veronica Fibisan,[3] Mark Dayer [iD] ,[5] Bernard D Prendergast [iD] ,[6] Peter Lockhart,[2] Larry Baddour [iD] ,[7] Jon Nicoll[3]

Case-crossover study

Thornhill M, Heart 2022

BDMA et Etude de type Case-Crossover

- **Toutes les admissions pour EI Angleterre** Avril 2010 - Mars 2016**: 14 731 sujets**

- **Procédures invasives dans les 15 mois qui précèdent**

- **Comparaison incidence**
  - Période Cas   3 mois avant admission pour EI
  - Période Contrôle   12 mois avant la période Cas

# BDMA et Etude de type Case-Crossover

- **PM / Défibrillateurs** OR 1.54 (95% CI 1.27 to 1.85)
- **Extraction dentaire** OR 2.14 (95% CI 1.22 to 3.76)
- **Gastroscopie haute** OR 1.58 (95% CI 1.34 to 1.85)
- **Gastroscopie basse** OR 1.66 (95% CI 1.35 to 2.04)
- **Biopsie médullaire** OR 1.76 (95% CI 1.16 to 2.69)
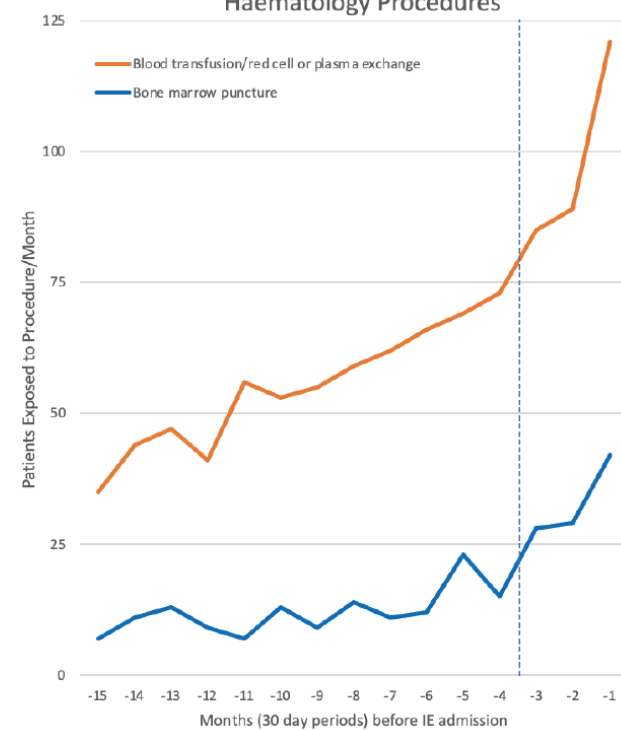- **Transfusion sanguine** OR 1.20 (95% CI 1.07 to 1.35)

Thornhill M, Heart 2022

BDMA et Etude de type Case-Crossover

Thornhill M, Heart 2022
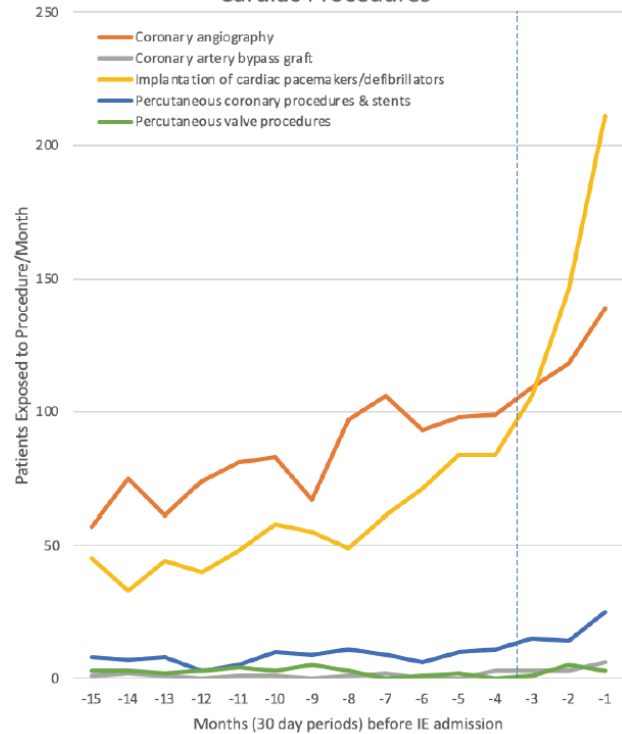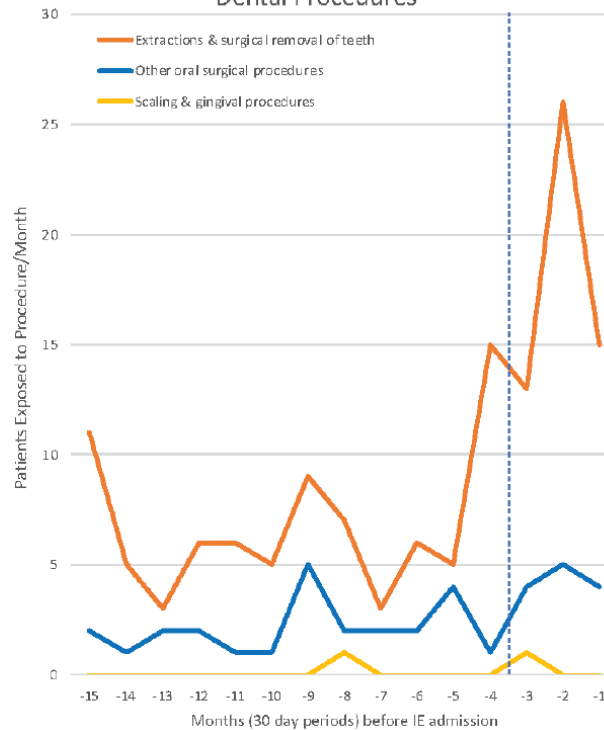
BDMA et Endocardite

Essais Randomisés

Essai randomisé en cluster sur registre (Registry-based cluster randomized trial)

Efficacité de l'antibioprophylaxie de l'endocardite infectieuse avant gestes dentaires invasifs chez les porteurs de prothèse valvulaire

- Dentistes d'un même cluster
- Patients porteurs de prothèse valvulaire
- Randomisation
- Système National des Données de Santé

PROPHETS

CHU BDX — CENTRE HOSPITALIER UNIVERSITAIRE BORDEAUX

Geste invasif

Clusters sans intervention auprès des dentistes

SNDS

Diagnostic endocardite bras **sans intervention**

Follow-up ≤ 3 mois

SNDS

Diagnostic endocardite bras **intervention**
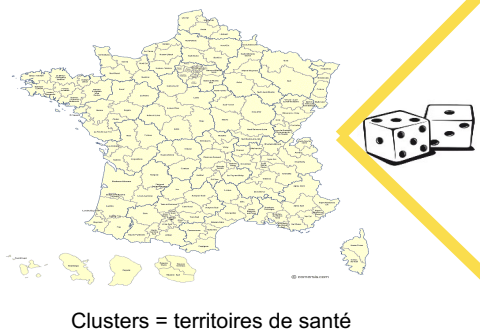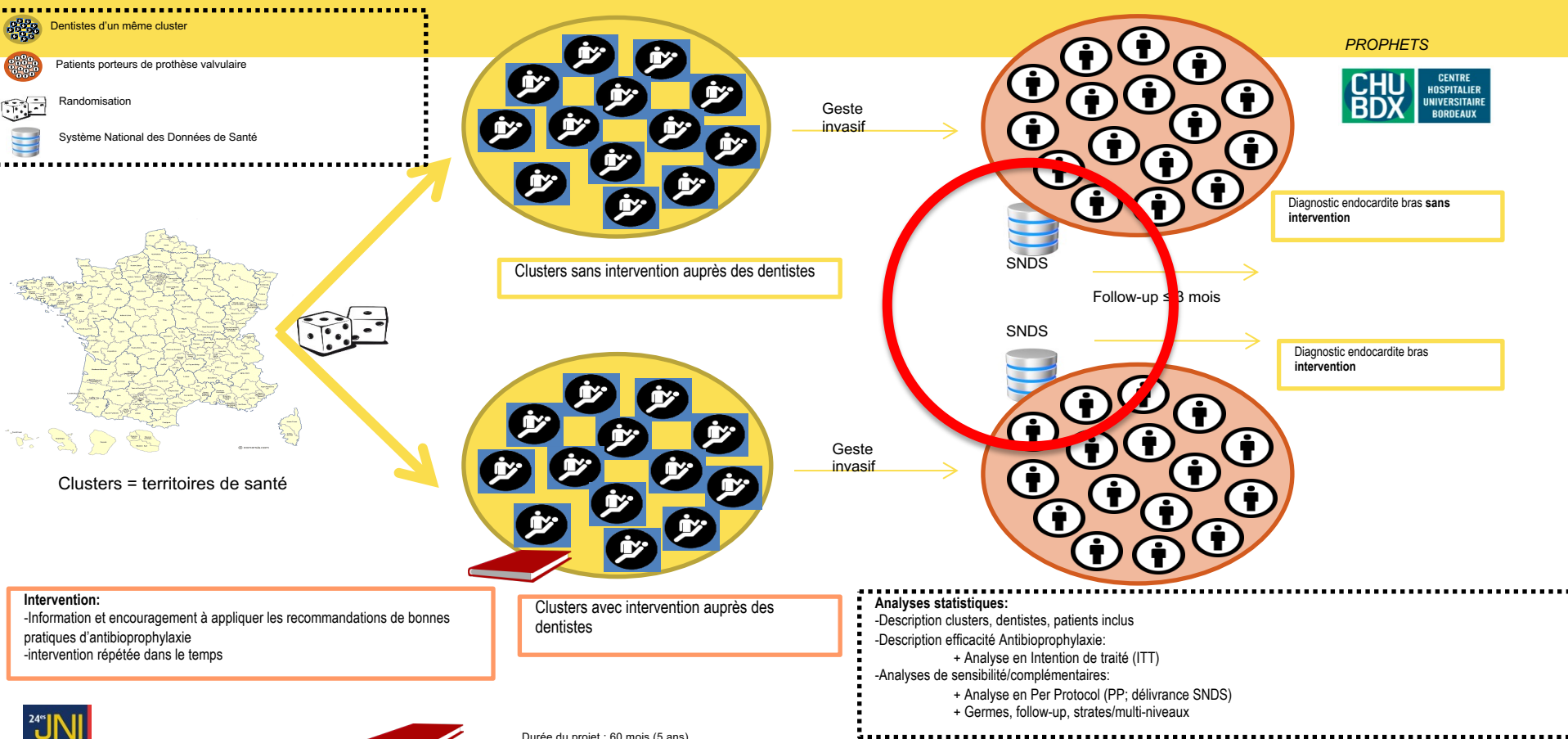
Clusters = territoires de santé

Geste invasif

Clusters avec intervention auprès des dentistes

Intervention:
-Information et encouragement à appliquer les recommandations de bonnes pratiques d'antibioprophylaxie
-intervention répétée dans le temps

Analyses statistiques:
-Description clusters, dentistes, patients inclus
-Description efficacité Antibioprophylaxie:
         + Analyse en Intention de traité (ITT)
-Analyses de sensibilité/complémentaires:
         + Analyse en Per Protocol (PP; délivrance SNDS)
         + Germes, follow-up, strates/multi-niveaux

24es JNI, GRENOBLE

Durée du projet : 60 mois (5 ans)
Budget: CHU 980 000€

AVANTAGES
Inconvénients

Avantages

- **Enregistrement passif**
- **Très grand nombre d'individus malgré maladie rare**

BDMA et Endocardite

- **Les données sont-elles valides ?**

- **Les analyses sont-elles valides ?**

BDMA et Endocardite

- **Les données sont-elles valides ?**
  - Les codages ICD identifient-t-ils correctement les EI ?

Les codages ICD identifient-t-ils correctement les EI ?

Fawcett *et al. BMC Medicine* (2019) 17:169
https://doi.org/10.1186/s12916-019-1390-x

Beyond Big Data to new Biomedical and Health Data Science: moving to next century precision health

BMC Medicine

RESEARCH ARTICLE

Open Access

Check for updates

'Caveat emptor': the cautionary tale of endocarditis and the potential pitfalls of clinical coding data—an electronic health records study

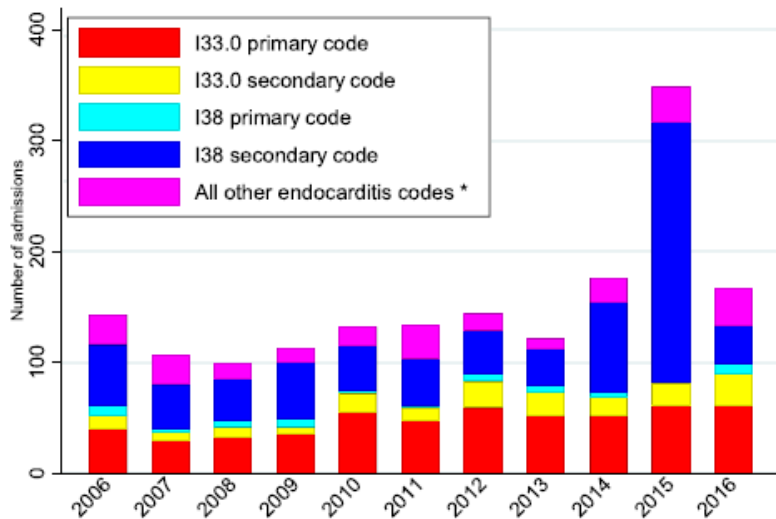Nicola Fawcett[1,2,3,9]*, Bernadette Young[2,3], Leon Peto[1,2,3], T. Phuong Quan[1,2,4], Richard Gillott[5], Jianhua Wu[6], Chris Middlemass[3], Sheila Weston[3], Derrick W. Crook[1,2,3,4], Tim E. A. Peto[1,2,3,4], Berit Muller-Pebody[7], Alan P. Johnson[1,7], A. Sarah Walker[1,2,4†] and Jonathan A. T. Sandoe[8†]

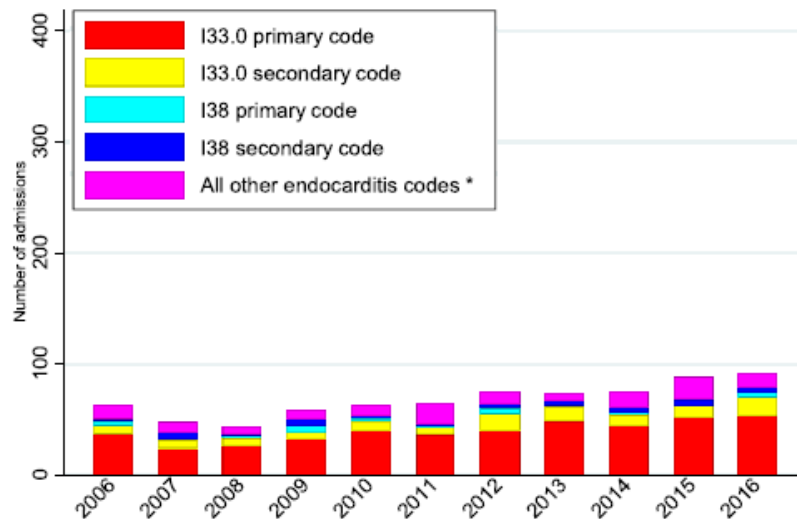# Les codages ICD identifient-t-ils correctement les EI ?

- **Deux Hôpitaux tertiaires Anglais 2006-2016 (Leeds, Oxford)**

- **Bases de données EI / Classification de Duke**

- **Leeds: ICD 1681 EI**

  - 738/1681**(44%)** des pts **ICD endocardite** : **EI certaine ou possible**

  - 263/1001**(24%)** de EI certaines ou possibles **NE** sont **PAS** codées EI

- "Estimating endocarditis incidence using raw admission data overestimated incidence trends twofold.

Total number of admissions with an endocarditis code

Leeds

Legend:
- I33.0 primary code
- I33.0 secondary code
- I38 primary code
- I38 secondary code
- All other endocarditis codes *

Number of admissions with an endocarditis code that represent a confirmed clinical case of endocarditis

Legend:
- I33.0 primary code
- I33.0 secondary code
- I38 primary code
- I38 secondary code
- All other endocarditis codes *

VPN et VPP des différents codes ICD différentes

Les codages ICD identifient-t-ils correctement les EI ?

- "Estimating endocarditis incidence using raw admission data **overestimated incidence trends twofold**".

- "Estimating incidence of streptococcal endocarditis using secondary codes also **overestimated** increases in incidence over time".

- Raisons:
  - Changement dans les pratiques de codage avec le temps
  - Changement dans les recommandations de codage

Electronic medical record-based deep data cleaning and phenotyping improve the diagnostic validity and mortality assessment of infective endocarditis: medical big data initiative of CMUH

Hsiu-Yin Chiang [a], Li-Ying Liang [b], Che-Chen Lin [a], Yi-Jin Chen [c], Min-Yen Wu [a], Sheng-Hsuan Chen [a], Pin-Hua Wu [d], Chin-Chi Kuo [a,c,e,f], Chih-Yu Chi [b,f],*

Les codages ICD identifient-t-ils correctement les EI ?

Table 2. Comparison of positive predictive value and age-adjusted in-hospital mortality according to different case identification strategies.

Case identification strategies	Sample size	Crude mortality (%)	Age-adjusted in-hospital mortality[a]
ICD	593	17.4	15.9

Les codages ICD identifient-t-ils correctement les EI ?

Table 2. Comparison of positive predictive value and age-adjusted in-hospital mortality according to different case identification strategies.

Case identification strategies	Sample size	PPV	Crude mortality (%)	Age-adjusted in-hospital mortality[a]
ICD	593	0.57 (0.53-0.61)	17.4	15.9
ICD and Duke-confirmed by chart review (Reference standard)[b]	336	-	24.4	21.0

ICD, International Classification of Diseases; PBC, positive blood culture; PPV, positive predictive value.

[a] Mortality was adjusted by age using logistic regression.

[b] Chart review was performed using the Duke criteria and definite or possible cases were considered.

Les codages ICD identifient-t-ils correctement les EI ?

Table 2. Comparison of positive predictive value and age-adjusted in-hospital mortality according to different case identification strategies.

Case identification strategies	Sample size	PPV	Crude mortality (%)	Age-adjusted in-hospital mortality[a]
ICD	593	0.57 (0.53-0.61)	17.4	15.9
ICD and (Fever or PBC or Vegetation)	373	0.78 (0.73-0.82)	20.9	19.4
ICD and (Fever or PBC)	368	0.76 (0.71-0.80)	21.7	19.8
ICD and (PBC or Vegetation)	363	0.90 (0.86-0.93)	24.5	21.8
ICD and (Fever or Vegetation)	347	0.81 (0.77-0.85)	21.6	19.9
ICD and Duke-confirmed by chart review (Reference standard)[b]	336	-	24.4	21.0
ICD and Vegetation	297	1.00 (0.99-1.00)	24.9	21.5
ICD and PBC	275	0.87 (0.82-0.90)	25.8	22.9
ICD and Fever	242	0.73 (0.67-0.79)	21.1	19.8
ICD and (PBC and Vegetation)	209	1.00 (0.98-1.00)	26.8	22.7
ICD and (Fever and PBC)	149	0.92 (0.86-0.96)	28.2	25.7
ICD and (Fever and Vegetation)	149	1.00 (0.98-1.00)	25.5	23.0
ICD and (Fever and PBC and Vegetation)	118	1.00 (0.97-1.00)	27.1	24.4

ICD, International Classification of Diseases; PBC, positive blood culture; PPV, positive predictive value.
[a] Mortality was adjusted by age using logistic regression.
[b] Chart review was performed using the Duke criteria and definite or possible cases were considered.

Les codages ICD identifient-t-ils correctement les EI ?

Supplementary Table 2. Demographic and clinical characteristics of patients with infective endocarditis confirmed on the basis of Duke criteria (definite or possible).

Variables	Patients with Duke-confirmed IE (N = 336)		P value
	With ICD of 424.9 or I38 N = 298 (88.7%)	Without ICD of 424.9 or I38 N = 38 (11.3%)	
Age (year, median [Q1, Q3])	59.1 (46.22, 72.16)	71.79 (52.46, 79.13)	
18-64 years	186 (62.42)	14 (36.84)	0.003
≥65 years	112 (37.58)	24 (63.16)	
Duke criteria			<0.0001
2 major	159 (53.36)	14 (36.84)	
1 major and 3-5 minor	46 (15.44)	5 (13.16)	
0 major and 5 minor	-	-	
1 major and 1-2 minor	85 (28.52)	11 (28.95)	
0 major and 3-4 minor	8 (2.68)	8 (21.05)	
0 major and 0-2 minor	-	-	
Valve replacement surgery [b]	53 (17.79)	4 (10.53)	0.26
Blood culture			
Two positive cultures within 14 days following IE diagnosis	216 (72.48)	22 (57.89)	0.06
Two positive cultures with typical pathogens [c]	176 (59.06)	20 (52.63)	0.45
Sonographic evidence of vegetation	273 (91.61)	24 (63.16)	<0.0001
Fever (≥ 38°C)	151 (50.67)	26 (68.42)	0.04

- **Les données sont-elles valides ?**
  - Les codages ICD identifient-t-ils correctement les EI ?
  - Les codages ICD identifient-ils correctement les porteurs de prothèse valvulaire cardiaque (PVC) ?

Les codages ICD identifient-t-ils correctement les PVC ?

- **Identification des 22 367 pts avec pose PVC en 2012 identifiés par un acte CCAM**

- **Recherche des séjours postérieurs à la pose pdt 10 ans**
  - Un des codages PMSI comprend la notion d'une PVC
    - 20 053 (89,6%) sont réhospitalisés
    - 16 194 (80,8%) sont identifiés comme PVC

Donc , au total, 28% des sujets NON identifiés dans les 10 ans qui suivent

- **Les données sont-elles valides ?**
- **Les analyses sont-elles valides ?**

Prevention

Figure 3: Incidence of infective endocarditis, by risk group

The figure shows the number of cases of infective endocarditis (superspells) recorded each month in individuals at high risk of developing infective endocarditis (solid red line) and those at lower risk (solid blue line). Data are corrected for change in the size of the total English population (not for change in the size of the high-risk or lower-risk groups). The vertical dashed line indicates March, 2008, the month in which cessation of antibiotic prophylaxis for infective endocarditis was recommended by the National Institute for Health and Clinical Excellence (NICE).[13] The trend lines for high-risk (dashed red line) and lower-risk (dashed blue line) individuals before and after introduction of the NICE guidelines are also shown.

Evidence that the upward trend in IE cases had changed
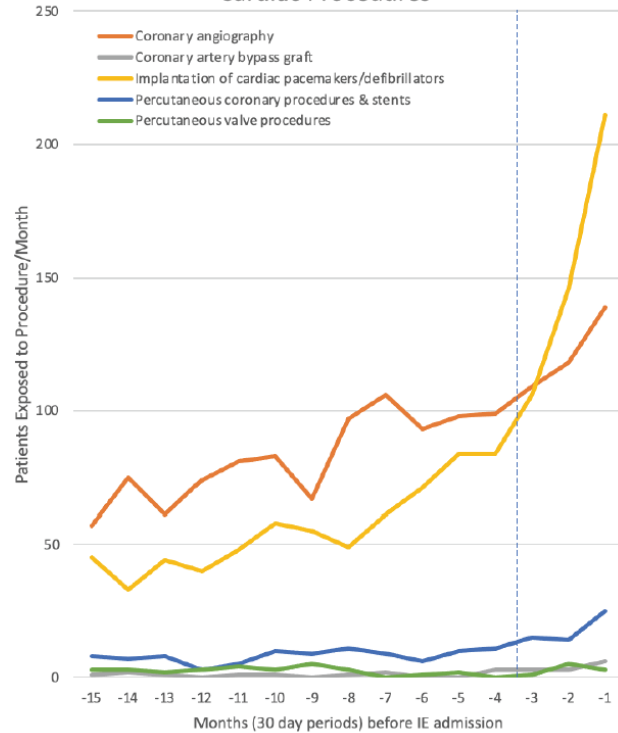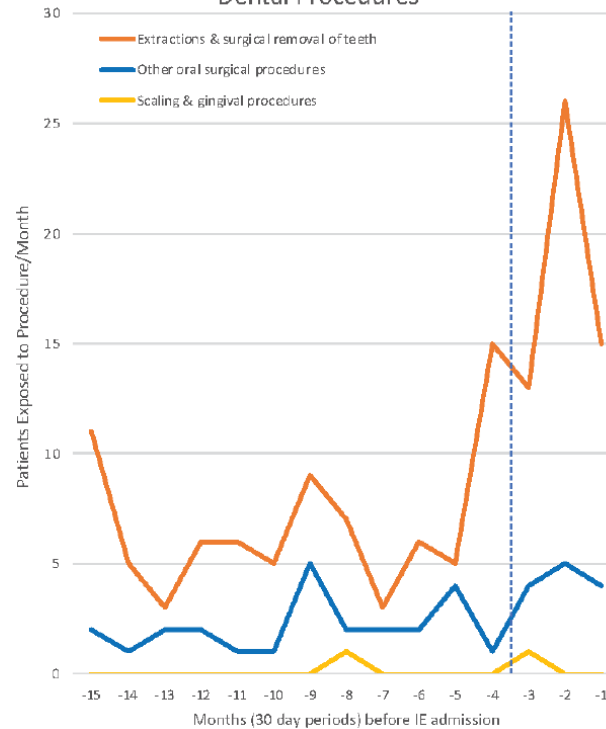**35 aditional IE cases each month**
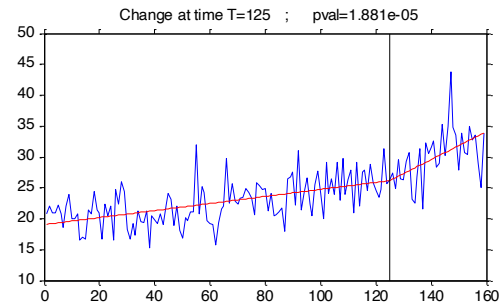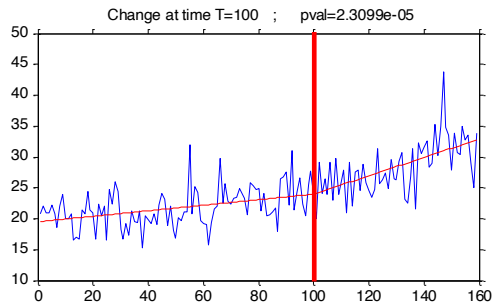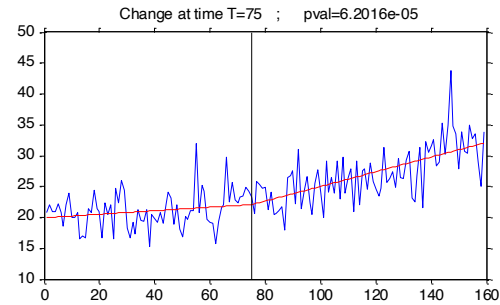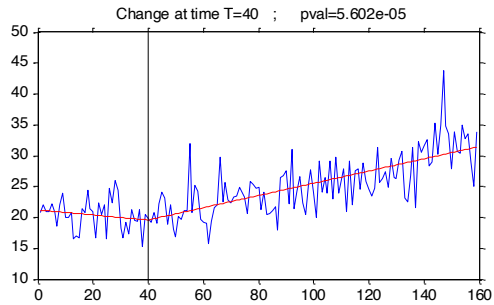
Figure 2 : 4 model fits under H1 (slope changes at time T), with T=40, 75, 100, 175
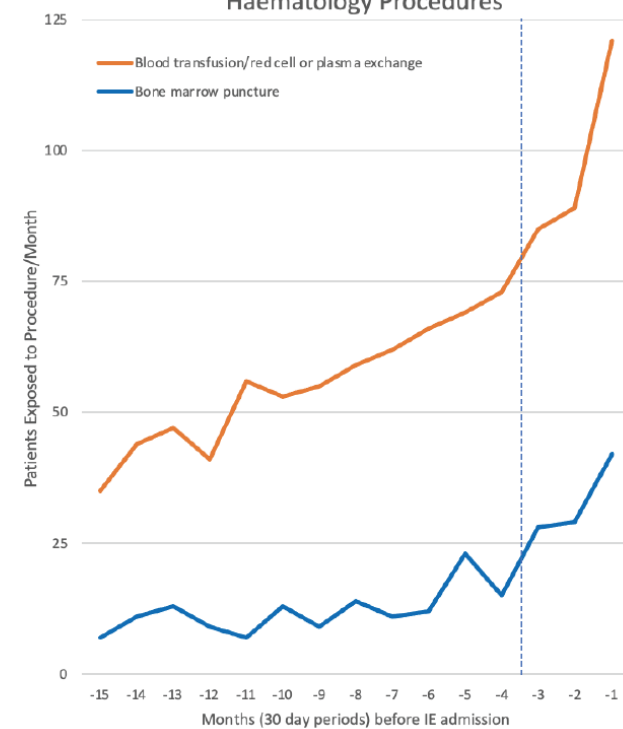Indeed, we see that we can reject H0 with a p-value < 0.0001 for any T between 40 and 130:

T=99 (March 2008).

B Iung Lancet 2015

BDMA et Etude de type Case-Crossover

Thornhill M, Heart 2022

Conclusions: BDMA et EI

- **BDMA: EDS hospitalières versus SNDS**
- **Qualité des données non établie**
- **Risque de biais différentiel**
- **Conséquences sur la validité des associations recherchées ?**
- **Association ne signifie pas causalité**
- **Analyses statistiques complexes nécessitant expertises**

Remerciements

- **Pr Vincent Le Moing**

- **Pr Bruno Hoen, Pr François Alla (PROPHETS)**

- **Dr Sarah Tubiana**

- **Dr Michaël Schwarzinger, Dr Jordan Guillot**