

# Epidémiologie prévisionnelle des maladies infectieuses

Antoine Flahault

Inserm - Université Pierre et Marie Curie UMR-S 707  
Ecole des Hautes Etudes en Santé Publique

21 mai 2007

UNIVERSITÉ  
PIERRE & MARIE CURIE  
LA SCIENCE À PARIS



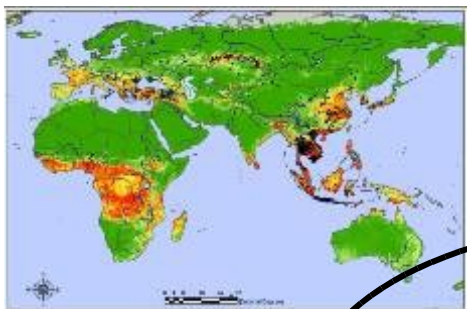
*WHO collaborating center for  
electronic disease surveillance*

**Inserm**

Institut national  
de la santé et de la recherche médicale

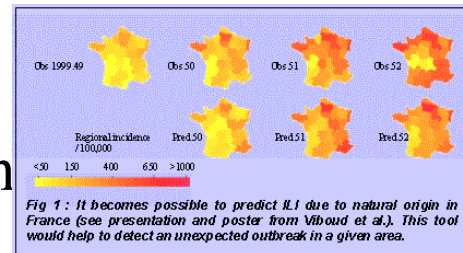
# Modèle = miroir simplifié de la réalité

## Modèle = « planche à dessin »



Réalité

validation



Observation

Compréhension

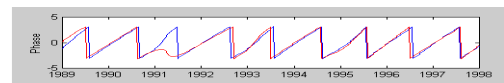
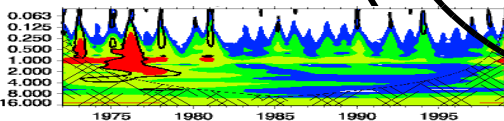
Détection

Prévision

calibration

Modèle

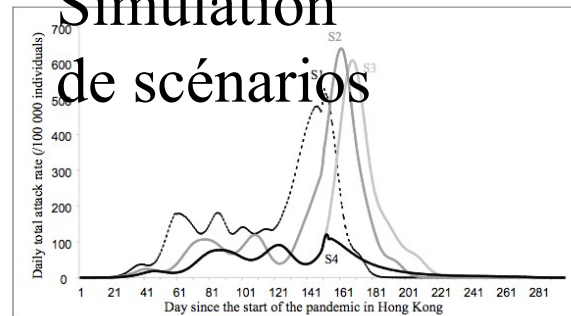
Simulation  
de scénarios



$$\frac{dX}{dt} = -\lambda X Y / N$$

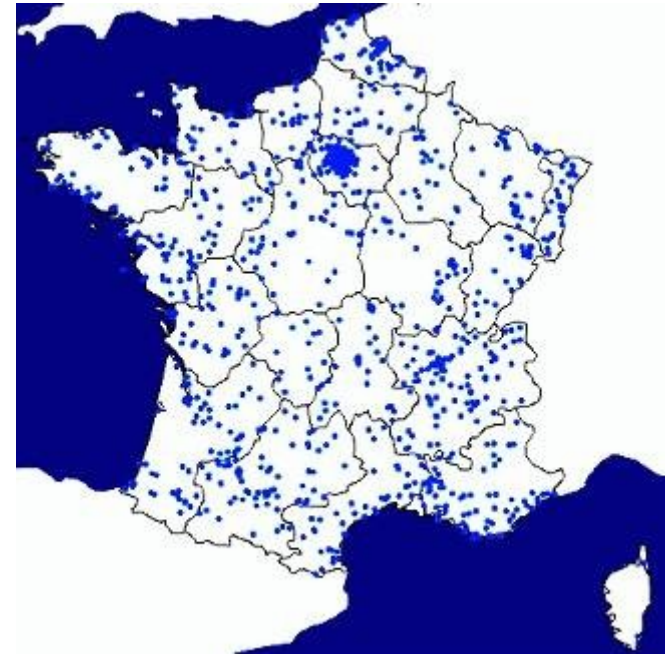
$$\frac{dY}{dt} = \lambda X Y / N - \gamma Y$$

$$\frac{dZ}{dt} = \gamma Y$$



# Le réseau *Sentinelles*<sup>®</sup>

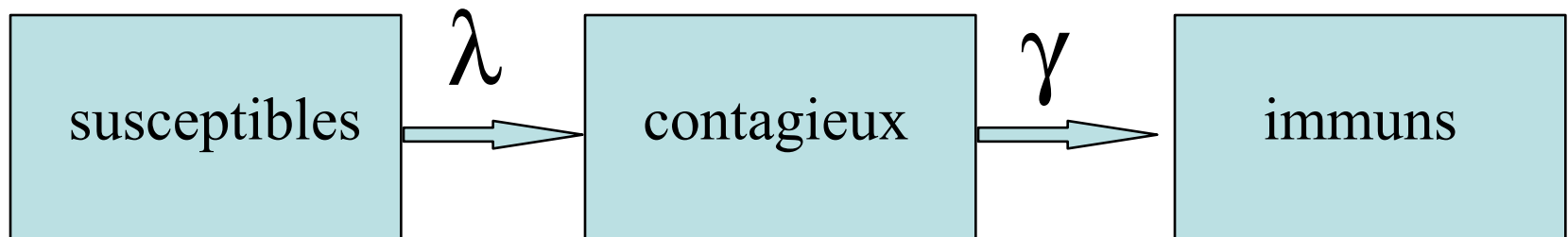
- 1193 médecins généralistes volontaires et bénévoles
- 14 indicateurs sous surveillance
  - Syndromes grippaux, diarrhées aiguës, Urétrites masculines, varicelle, zona, rougeole, oreillons, hépatites A-B-C, crises d'asthme, tentatives de suicides, hospitalisations
- Saisie électronique des données
- **Temps-réel**



*Valleron, Am J Publ Health, 1986*  
*Flahault, BMJ, 1997*

# Théorie mathématique des épidémies

un modèle compartimental simple



*1927 : Kermack & McKendrick*

# Le théorème du seuil

$$\frac{dY}{dt} = \beta c X Y / N - (1/d) \cdot Y > 0 \Rightarrow \text{épidémie}$$

$\beta$

*probabilité de transmission*

$c$

*nombre de contacts par unité de t*

$d$

*durée de la période contagieuse*

$$R_0 = \beta c d > 1$$

*taux de reproduction de base*

$$T_d = \frac{d \cdot \ln(2)}{R_0 - 1}$$

*temps de doublement*

# Ro : Détection des épidémies

Exemple de la grippe (réseau Sentinelles, France)

$$d = 4 \text{ j}$$

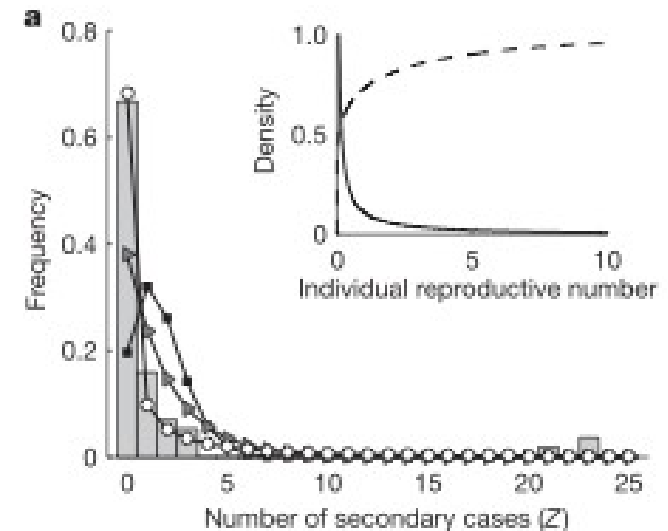
$$T_d = \frac{d \cdot \ln(2)}{R_0 - 1} = 3 \text{ j} \Rightarrow R_0 = \frac{d \cdot \ln(2) + T_d}{T_d} \approx 2$$

Doublement d'incidence en trois jours =>

$R_0 > 1 \Leftrightarrow$  épidémie sur le territoire national

# $R_0$ : « Une échelle de Richter » pour les maladies transmissibles ?

Rougeole	$R_0 = 15$ à $20$
Grippe	$R_0 = 1,8$ à $2,5$
Variolle	$R_0 = 3$
SRAS	$R_0 = 2$
Hépatite B	
-Groupes à haut risque	$R_0 = 4$ à $8,8$
-Population générale	$R_0 = 1,1$



*$R_0$  du SRAS à Singapore (Lloyd-Smith, 2005)*

# Application du théorème du seuil stratégies vaccinales et immunité grégaire

- quelle proportion de la population faut-il immuniser pour bloquer le déclenchement d'une épidémie ?

$$p > (1 - 1/R_0)$$

Pour la rougeole	(R0 = 15-20)	$p = 93-95\%$
Pour la grippe	(R0 = 2-4)	$p = 50-75\%$
Pour l'hépatite B		
- chez les ind. à ht risque	(R0 = 4)	$p = 75\%$
- chez les ind. à faible risque	(R0 = 1,1)	$p = 10\%$
- Milieux à très haut risque	(R0 = 8,8)	$p = 89\%$

# Application du théorème du seuil calendrier vaccinal

– quel est l'âge optimal pour vacciner ?

$$T = \frac{[\ln(1/D) - \ln(1/A)]}{[1/D - 1/A]}$$

$D$  durée de la protection par les Ac maternels

$A$  âge moyen de la maladie

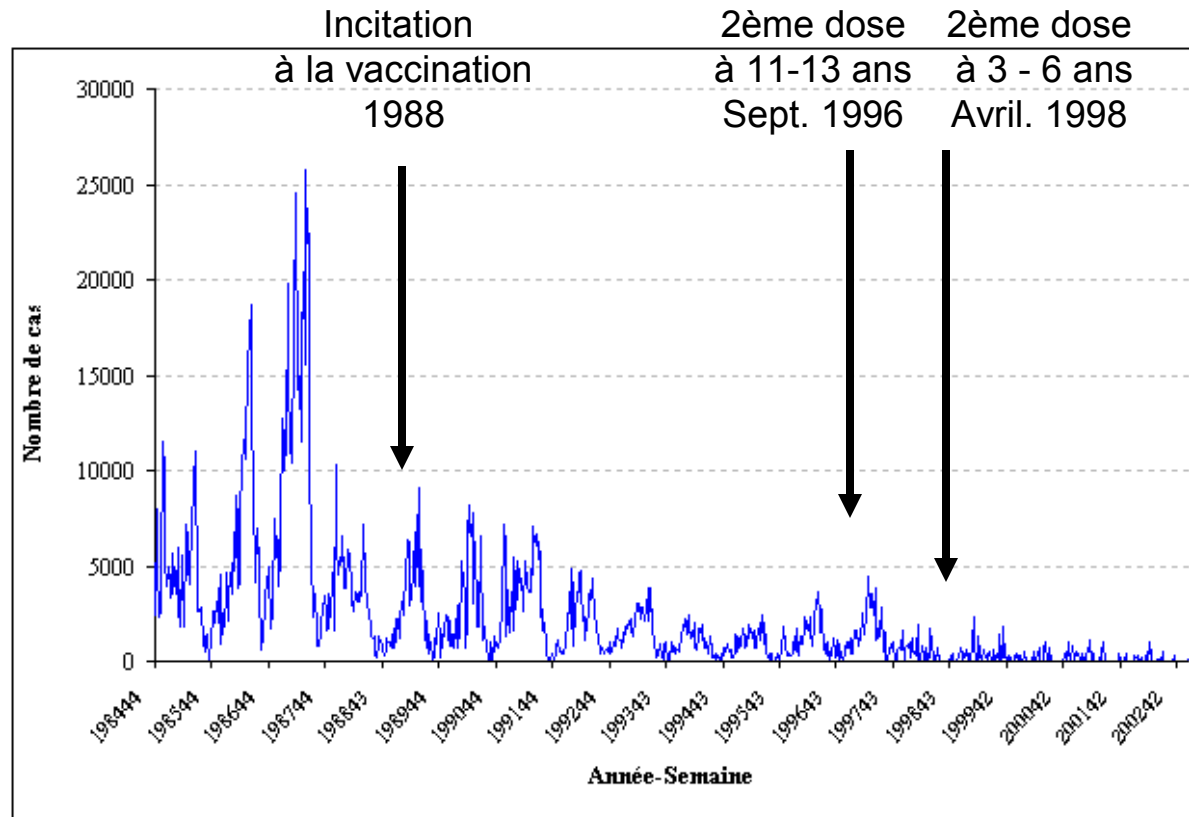
exemple : Rougeole en PVD,  $D=6$  mois,  $A=18$  mois,  
d'où  $T = 10$  mois

(Katzmann & Dietz, 1984)

depuis 1927 : Kermack & McKendrick

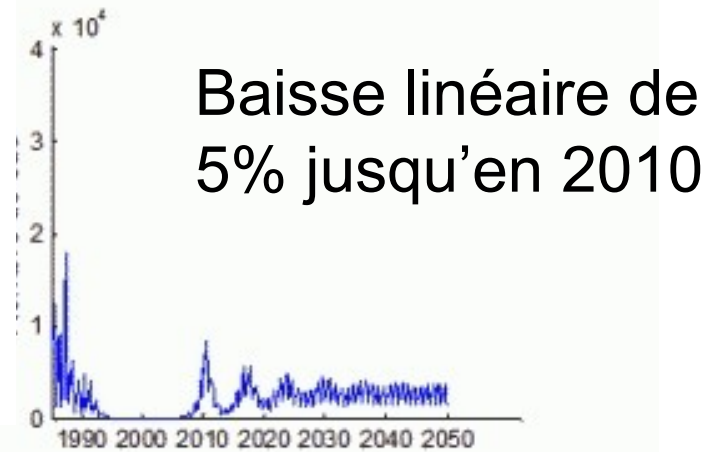
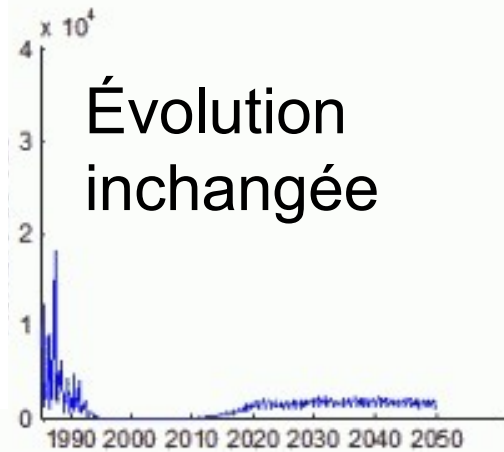
- rôle central dans la théorie mathématique des épidémies (grippe, SIDA),
- implications importantes sur les mises en place de stratégies vaccinales de masse (rougeole, oreillons, rubéole, coqueluche, varicelle)

# Observation : La rougeole en France 1984 - 2004



Source : réseau Sentinelles, Inserm

# Modélisation : Scénarios de baisse de la couverture vaccinale



(H. Sarter, 2004)

# Contrôle et prévention d'une future pandémie de grippe

Taux de reproduction de base

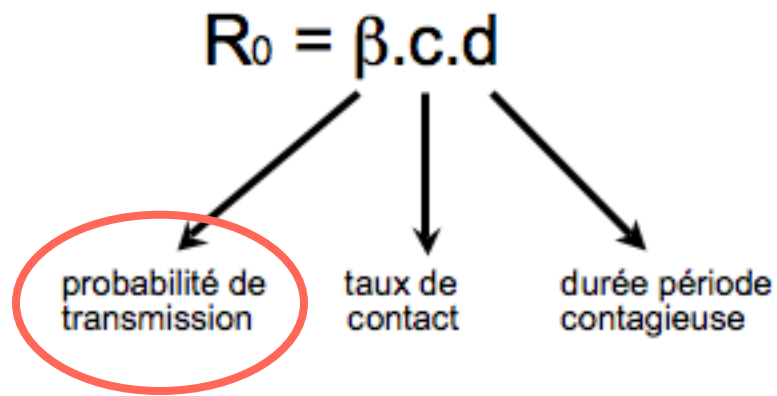
$$R_0 = \beta \cdot c \cdot d$$

probabilité de  
transmission

taux de durée période  
contact

contagieuse

Théorème du seuil, pandémie :  $R_0 > 1$



Antiviraux (curatifs, préventifs)

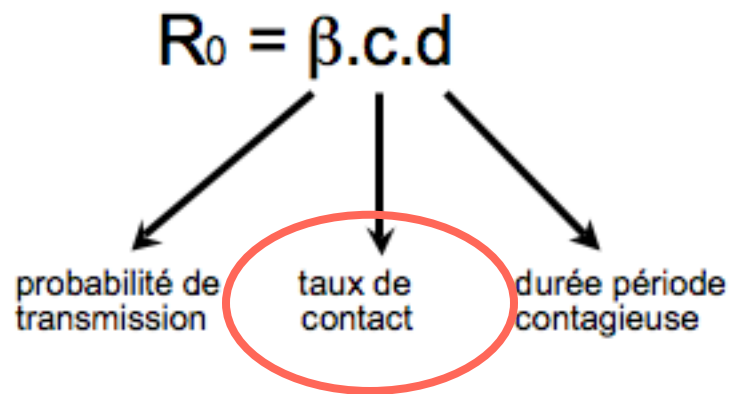
Résistances, associations d'antiviraux

Vaccins

Masques de protection

Lavage des mains

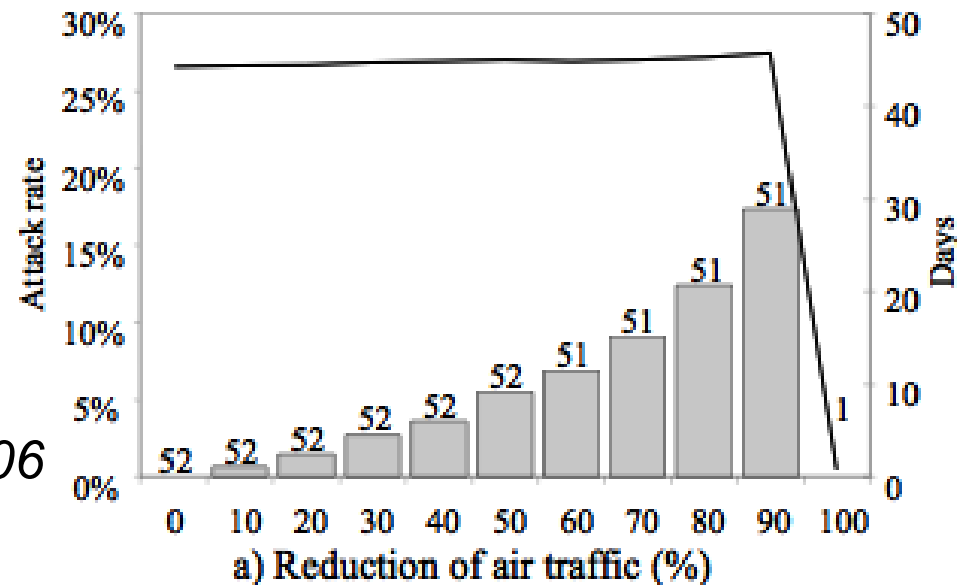
- vers un concept de recherche clinique rapide et réactive ?



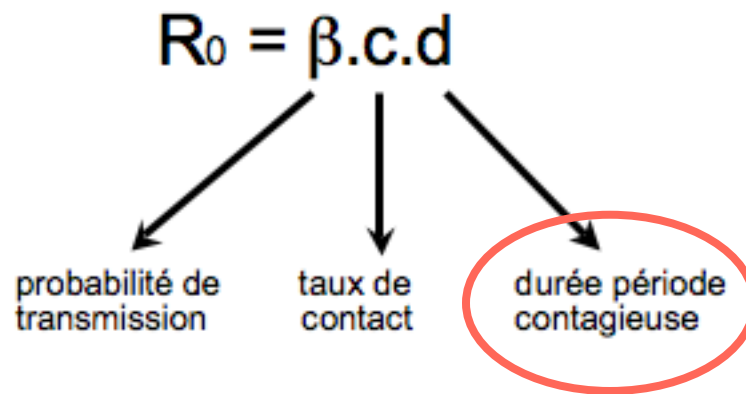
## Augmenter la « distance sociale »

- Isolement des malades
- Fermeture des crèches, écoles, universités
- Réduction des mouvements de population

Fig. 2.

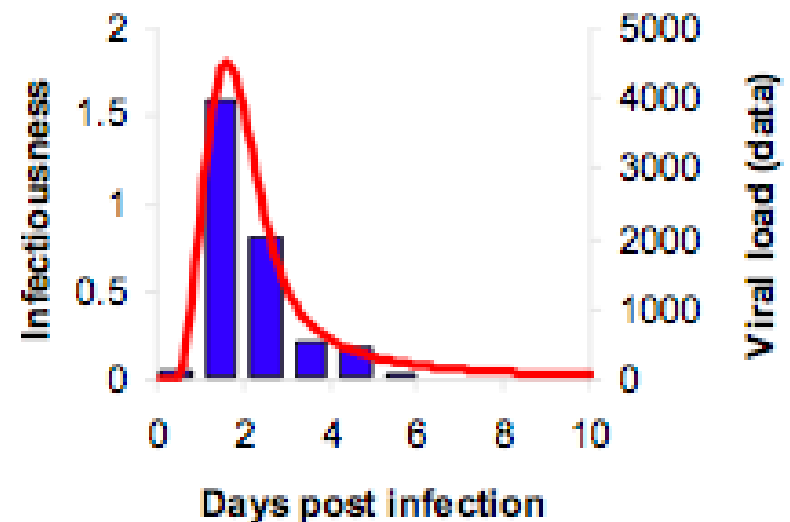


*Flahault A et coll, Vaccine 2006*



## Diminuer la durée de la période contagieuse

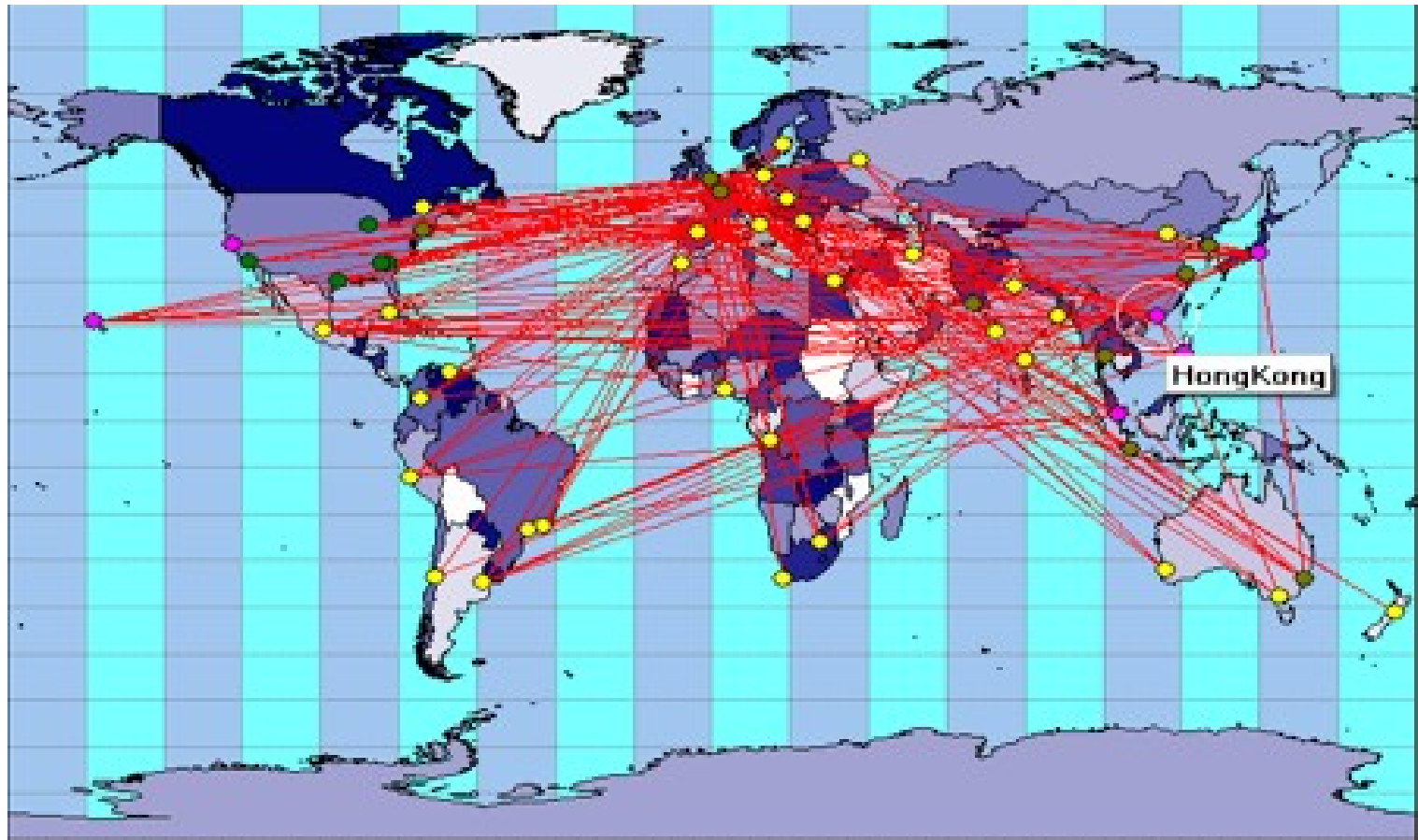
- Antiviraux
- Médicaments à visée symptomatique
- Mais déjà : connaître la durée de la période contagieuse : 4 -> 2,6 j



*Cauchemez S, Stat Med, 2004*

*Ferguson N, Nature 2005*

Simulation sur ordinateur d'une nouvelle pandémie  
*(réseau de 52 métropoles, données de transport aérien 2000)*



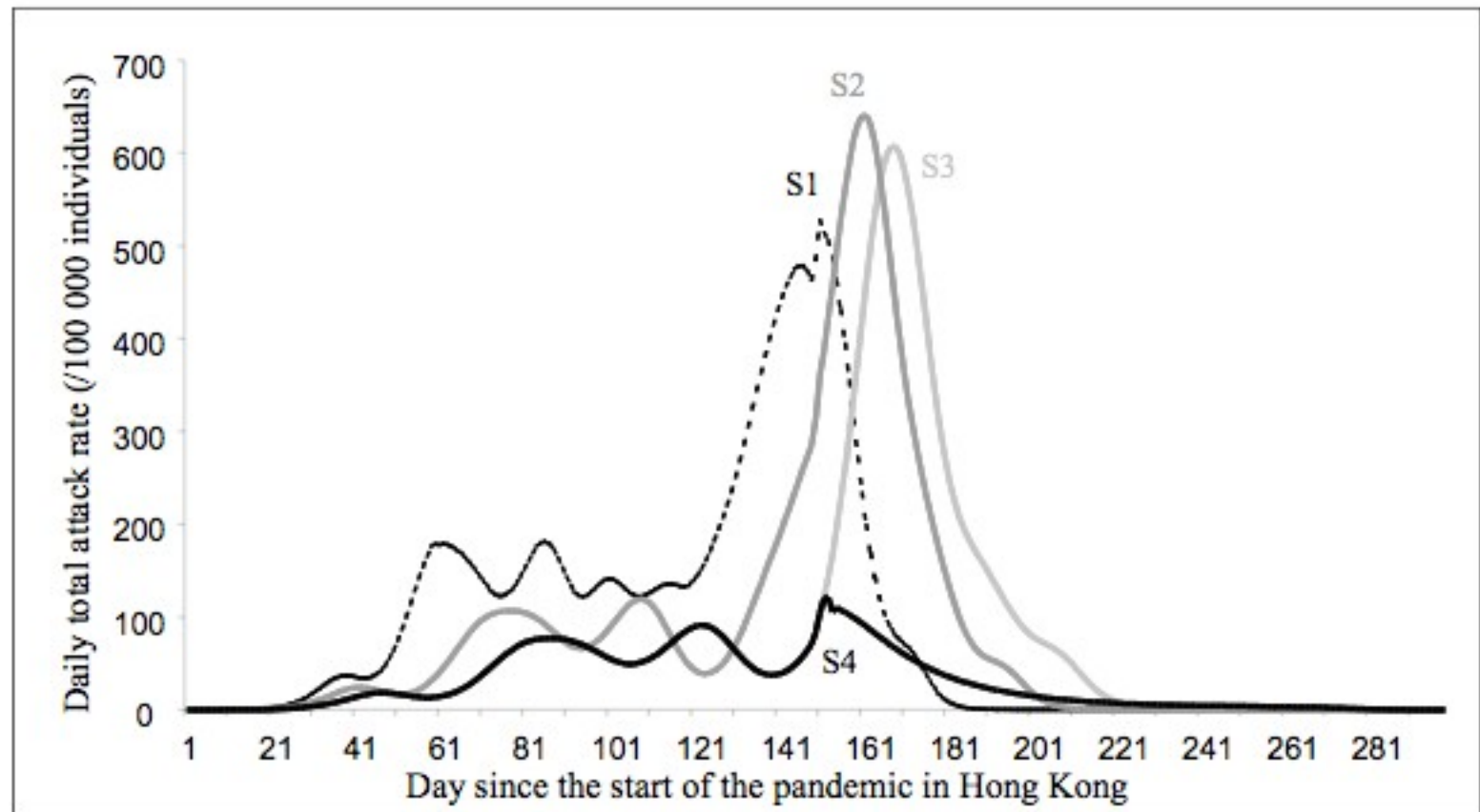
# Simulation de scénarios de prévention

*S1*: Pas de mesures de contrôle

*S2*: Isolement de 10% infect. et 60% restriction transport aérien dès J1

*S3*: comme *S2* + 5 jours d'antiviraux pour 60% des syndromes grippaux.

*S4*: comme *S3* + vaccination 70% efficace. chez 50% susceptibles dès J151



# Modèles individus centrés

Vol 437|8 September 2005|doi:10.1038/nature04017

nature

ARTICLES

---

## Strategies for containing an emerging influenza pandemic in Southeast Asia

Neil M. Ferguson<sup>1,2</sup>, Derek A.T. Cummings<sup>3</sup>, Simon Cauchemez<sup>4</sup>, Christophe Fraser<sup>1</sup>, Steven Riley<sup>5</sup>, Aronrag Meeyai<sup>1</sup>, Sophon Iamsrithaworn<sup>6</sup> & Donald S. Burke<sup>3</sup>

Highly pathogenic H5N1 influenza A viruses are now endemic in avian populations in Southeast Asia, and human cases continue to accumulate. Although currently incapable of sustained human-to-human transmission, H5N1 represents a serious pandemic threat owing to the risk of a mutation or reassortment generating a virus with increased transmissibility. Identifying public health interventions that might be able to halt a pandemic in its earliest stages is therefore a priority. Here we use a simulation model of influenza transmission in Southeast Asia to evaluate the potential effectiveness of targeted mass prophylactic use of antiviral drugs as a containment strategy. Other interventions aimed at reducing population contact rates are also examined as reinforcements to an antiviral-based containment policy. We show that elimination of a nascent pandemic may be feasible using a combination of geographically targeted prophylaxis and social distancing measures, if the basic reproduction number of the new virus is below 1.8. We predict that a stockpile of 3 million courses of antiviral drugs should be sufficient for elimination. Policy effectiveness depends critically on how quickly clinical cases are diagnosed and the speed with which antiviral drugs can be distributed.

## Expected pattern of spread of an uncontrolled epidemic $R_0=1.5$ .

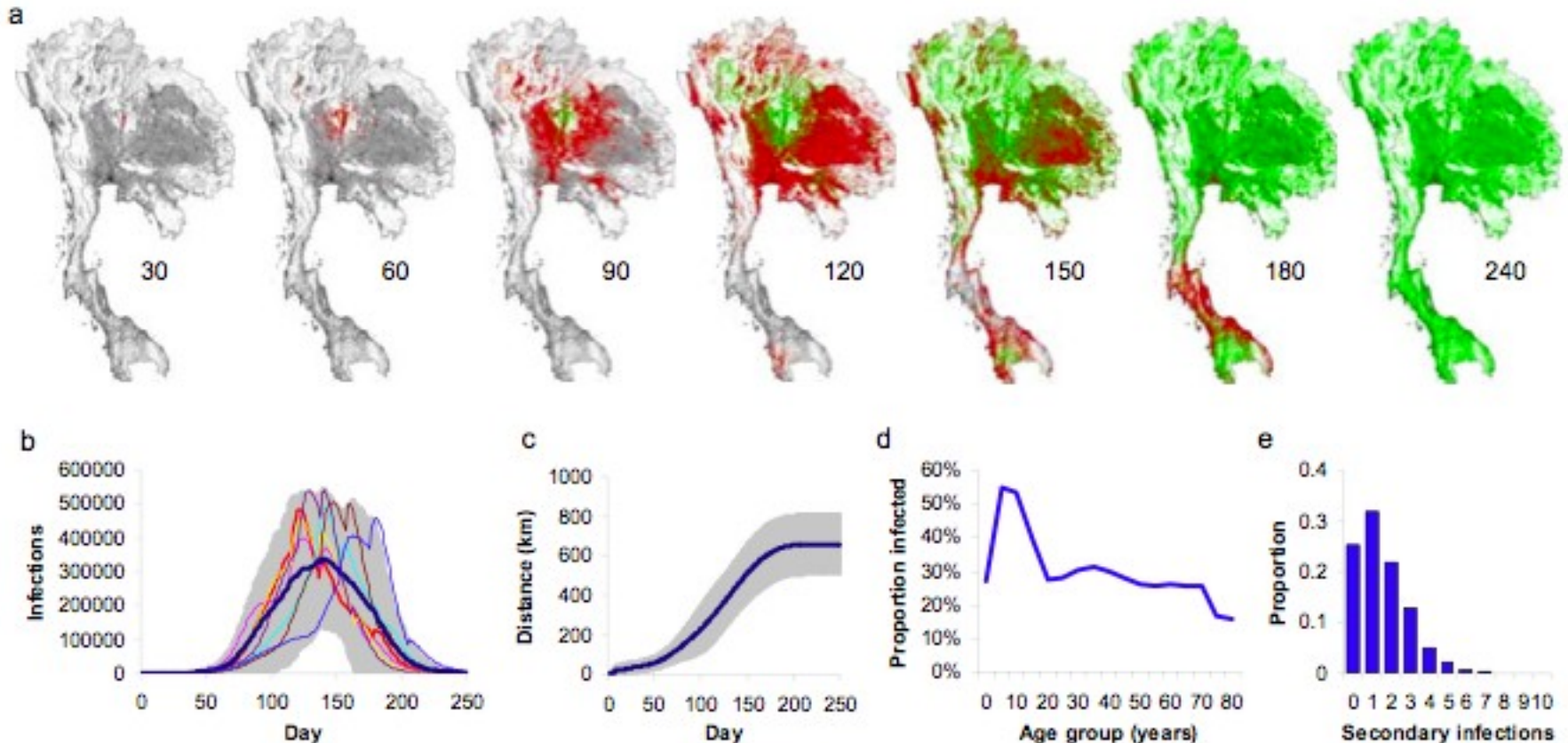
(a) Spread of a single simulation. Red = infectives, green = recovered from infection or died.

(b) Daily incidence of infection. Thick blue line = average, grey shading = 95% envelope of incidence timeseries. Multiple coloured = a sample of realisations.

(c) Root Mean Square (RMS) distance from seed infective of all individuals infected since the start of the epidemic as a function of time.

(d) Attack rate by age (mean = 33%).

(e) Number of secondary cases per primary case

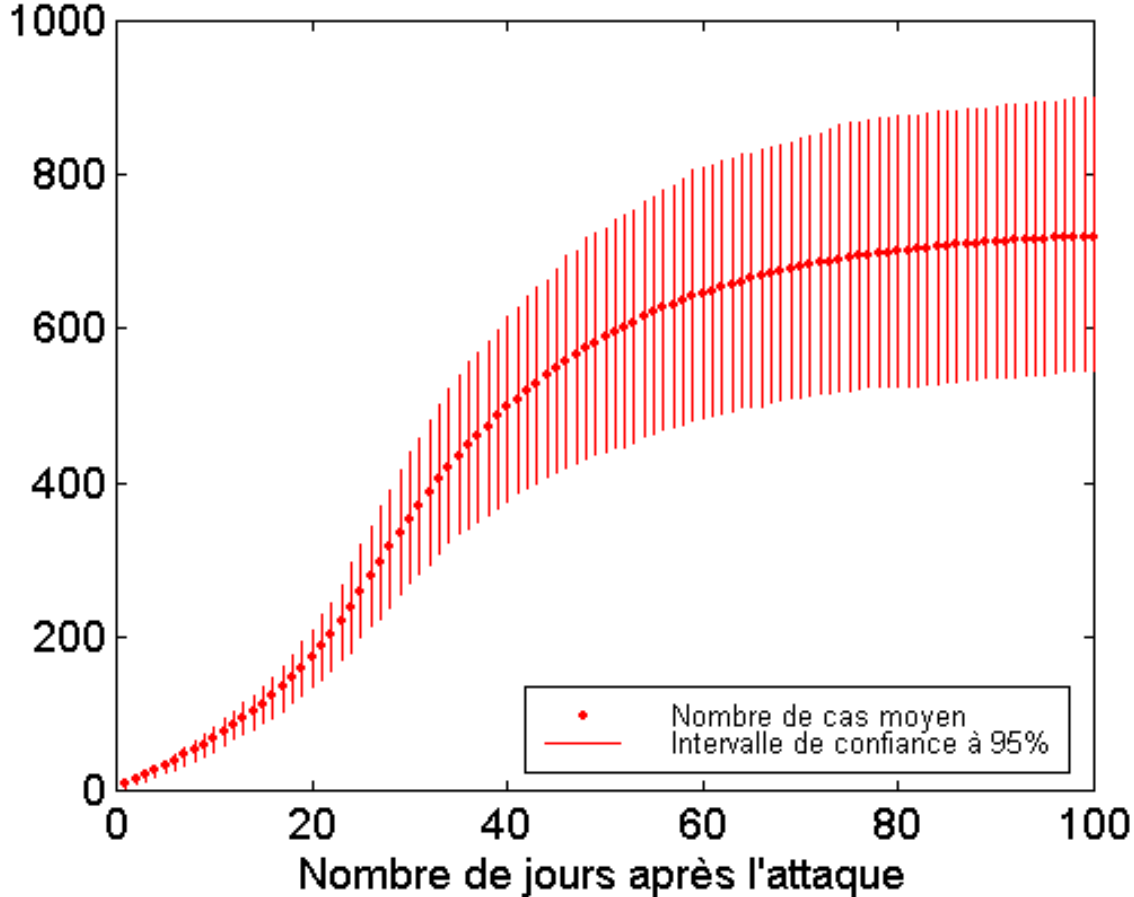


# **Le travail de Ferguson et coll. : la plus grande simulation sur ordinateur jamais publiée**

- Simulation d'une population de 85 millions de personnes en Thaïlande
- 10 ordinateurs haute capacité en parallèle
- > 1 mois de temps d'ordinateur

# Variole (bioterroriste) : modélisation

Nombre de cas cumulés



durée

105 [190;310] jours

cas vaccinaux

440 [3 910;6 840]

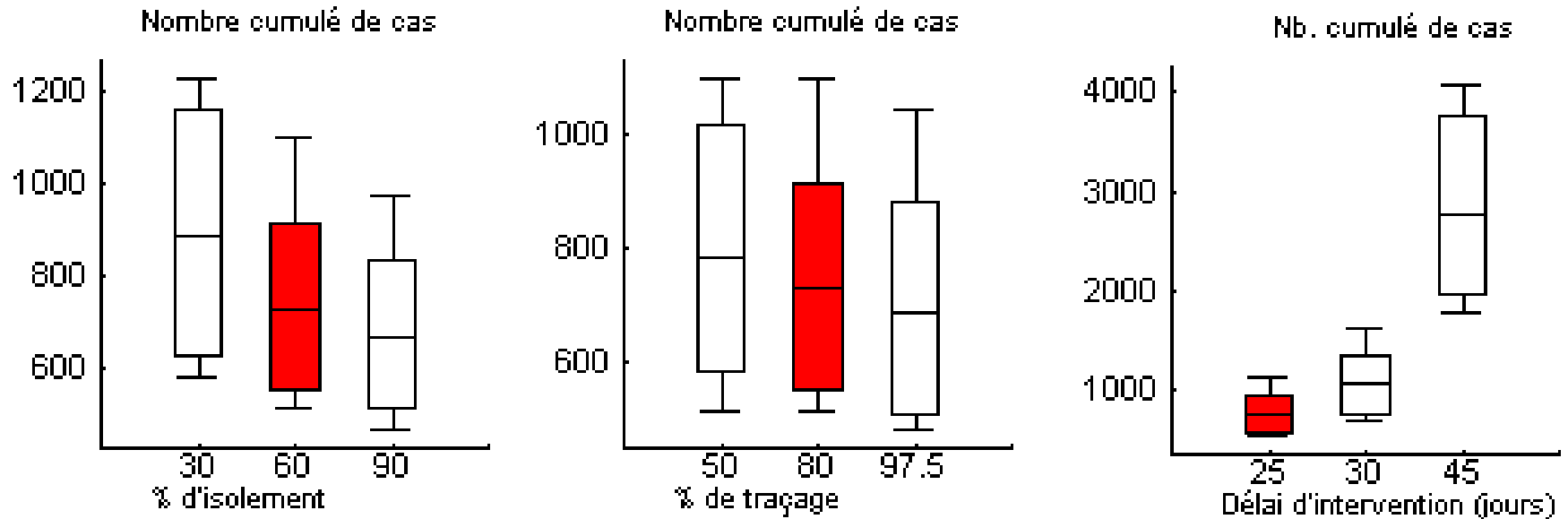
personnes isolées

50 [415;686]

max. de personnes isolées

50 [170;300]

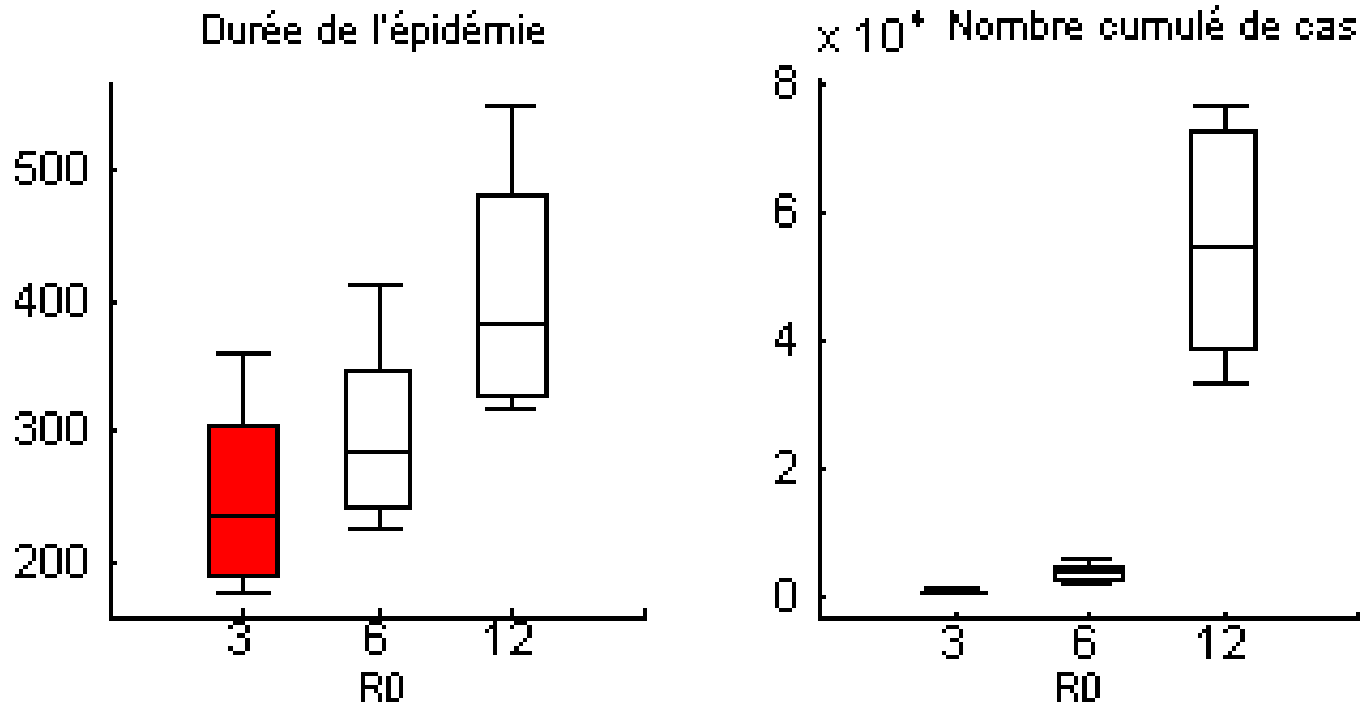
# Importance du délai d'intervention *(sur la taille de l'épidémie)*



 Scénario de référence

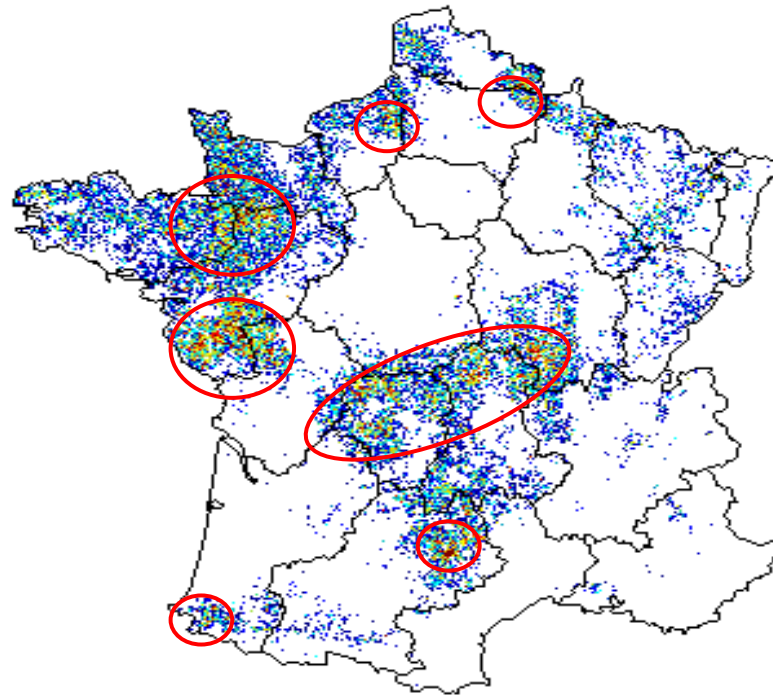
# Varirole : Influence du $R_0$

- Débat sur la valeur du  $R_0$  (*Enserink, Science, 2002*)



# Fièvre aphteuse : identification des zones à risque

Carte du *Ro*



# Maladies émergentes, environnement et technologies spatiales

UNIVERSITÉ  
PIERRE & MARIE CURIE  
LA SCIENCE À PARIS



*WHO collaborating center for  
electronic disease surveillance*

**Inserm**

Institut national  
de la santé et de la recherche médicale

# Prospective 1920...

**Le plus efficace outil d'alerte précoce des épidémies de paludisme (et peut-être encore de toutes les maladies vectorielles) date de 1923 !**

Une combinaison de :

- l'absence de paludisme dans les 5 dernières années
- pluies d'intensité anormale entre juillet et août
- prix local des céréales (reflet du statut nutritionnel de la population)

Utilisé pendant 25 ans avec une précision remarquable, puis abandon en 1950 (arrivée des traitements, mais apparition de résistances...)

*d'après Gill C.A. The prediction of malaria epidemic. Ind J Med Res 10, 1136-43 (1923)*

# ...Prospective 2000...

- Plutôt que jauger le niveau de précipitation et des prix des céréales
- De nombreuses informations ont été trouvées associées à la dynamique des populations de vecteurs et de la transmission des pathogènes
  - facteurs environnementaux
  - données satellitaires

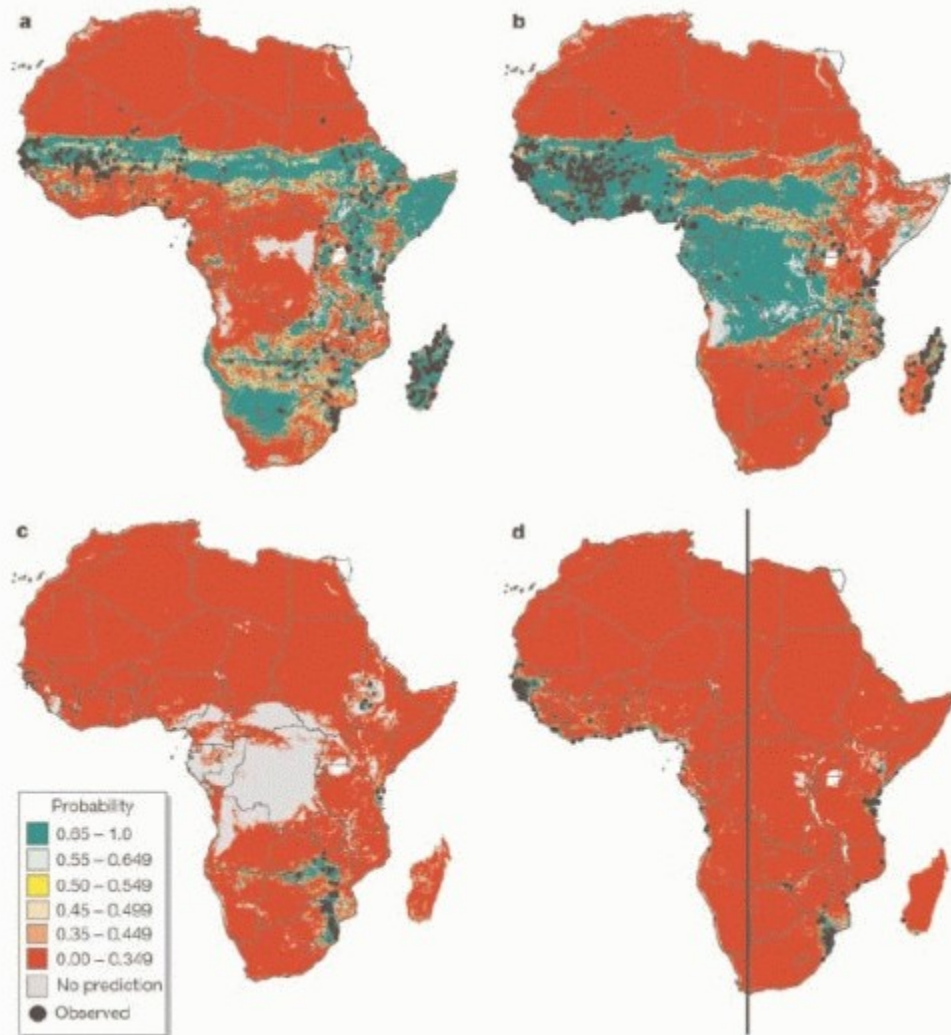
# Données satellitaires

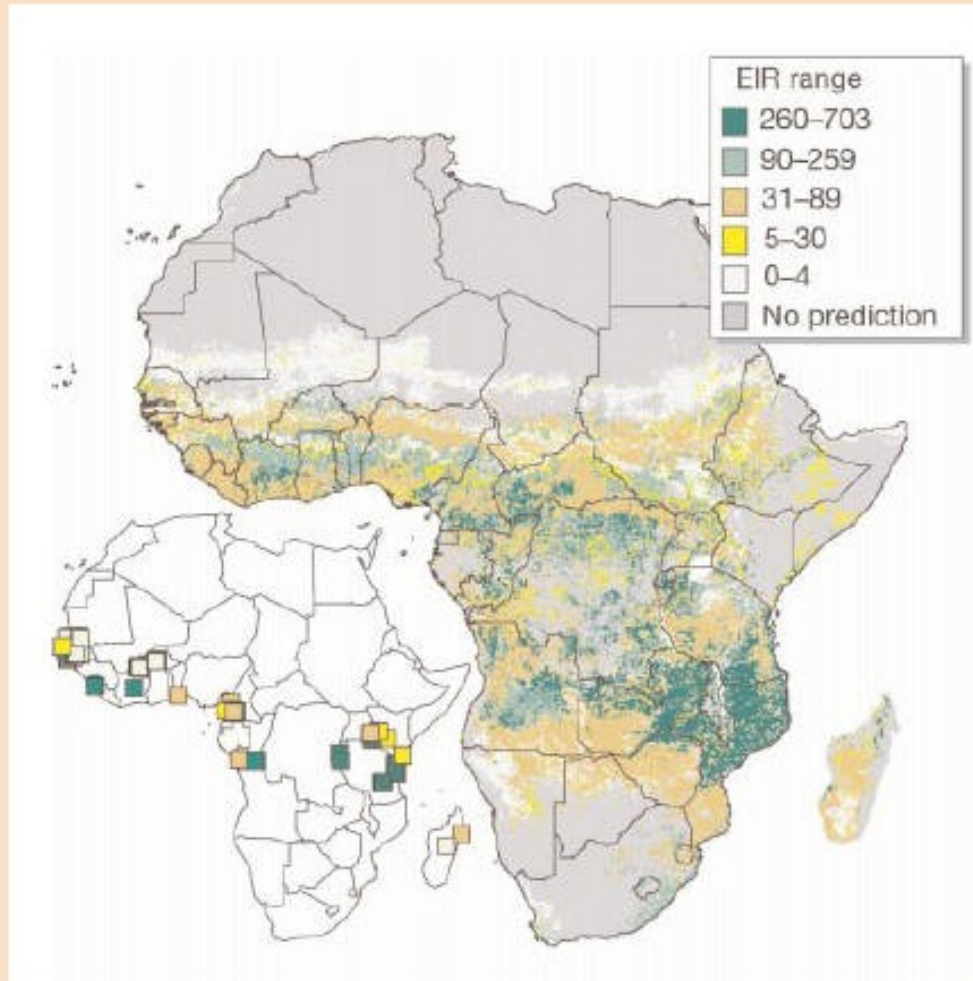
- Rarement idéales pour l'épidémiologie
  - compromis entre la résolution spectrale, spatiale et temporelle, contraintes
  - données de capture passive surtout, mais intérêt croissant pour les images radar
- Résolution spectrale
  - spectre visible jusqu'à l'infrarouge
- Résolution spatiale
  - Ikonos, SPOT, Landsat : 1-120 m
  - Meteosat, NOAA-AVHRR : 1-8 km
- Résolution temporelle
  - Ikonos (11j), SPOT (26j), Landsat (16j) => épidémiol. analytique
  - Meteosat : 2x/j à 2x/h (sat. géostat.)=> épidémiol. dynamique

# Du satellite à l'épidémiologie

- Les données brutes
  - température de surface des sols (LST)
  - index d'anomalie de végétation (NDVI)
  - température de surface de l'air
  - pluviométrie
- Modélisation et applications
  - Transformation de Fourier, ACP, moy.mobile => filtrer les séries
  - utilisation des séries transformées dans des modèles de régression logistique, ou des analyses discriminantes
  - Production de cartes de risque dans des lieux où la situation épidémiologique est bien documentée (jeu d'apprentissage)
    - prédictibilité d'un pixel vis-à-vis de la présence du vecteur ou de la maladie
    - prédiction de l'impact de la maladie avec des données quantitatives
  - Evaluation de la qualité des prévisions obtenues

**Figure 2** Distributions of five mosquito species in the *Anopheles gambiae* complex in Africa, predicted from temporal Fourier-processed satellite data (Box 1) and elevation (global coverage provided by the digital elevation model GTOPO30; <http://edcdaac.usgs.gov/gtopo30/README.html>) at a spatial resolution of 0.05°. The colour-coded probabilities<sup>42</sup> of presence effectively indicate the environmental suitability for each species throughout the continent. Symbols indicate sample sites for each species. Between 18° N to 30° S, each species was classed as present only within 0.15° of the sites from which it has been recorded<sup>24</sup> and absent only from similarly sized sites where any of the other species have been recorded. Within these sites, presence pixels (300 for *A. gambiae* s.s. and *A. arabiensis* and 100 for the other species) and absence pixels (400 for all species) were chosen at random; additional randomly selected absence pixels were chosen north of 18° N ( $n = 200$ ) and south of 30° S ( $n = 50$ ). Satellite data: middle infrared, land surface temperature and normalized difference vegetation index for 1982–1998 were derived from the National Oceanographic and Atmospheric Administration's Advanced Very High Resolution Radiometer, and cold cloud duration for 1988–1999 from Meteosat High Resolution Radiometer. Satellite and elevation data for all sample pixels were subjected to *k*-means clustering within the Statistics Package for the Social Sciences (SPSS, Chicago) to identify up to six natural clusters each of presence and absence pixels for each mosquito species. Within maximum





**Figure 3** Satellite-derived predictions of entomological inoculation rate (EIR) in Africa. EIR data<sup>27</sup> (map inset) were grouped into five approximately equal-sized classes of mean levels of malaria challenge. The same satellite data layers and analytical methods as used in Fig. 2 are used to define the probability with which each continental pixel belongs to one of the five challenge categories. Insufficient training data were available to define EIR in those parts of the continent marked grey. Descriptive accuracies: EIR range 0–4.4, 85.7% ( $n = 21$ ); 5.8–26.6, 81.8% ( $n = 22$ ); 31.0–87.0, 77.3% ( $n = 22$ ); 89.8–255.5, 68.2% ( $n = 22$ ), 259.9–703.4, 95.5% ( $n = 22$ ). Overall  $\kappa = 0.771 (\pm 0.064)$ .

# L'actuelle "pandémie"

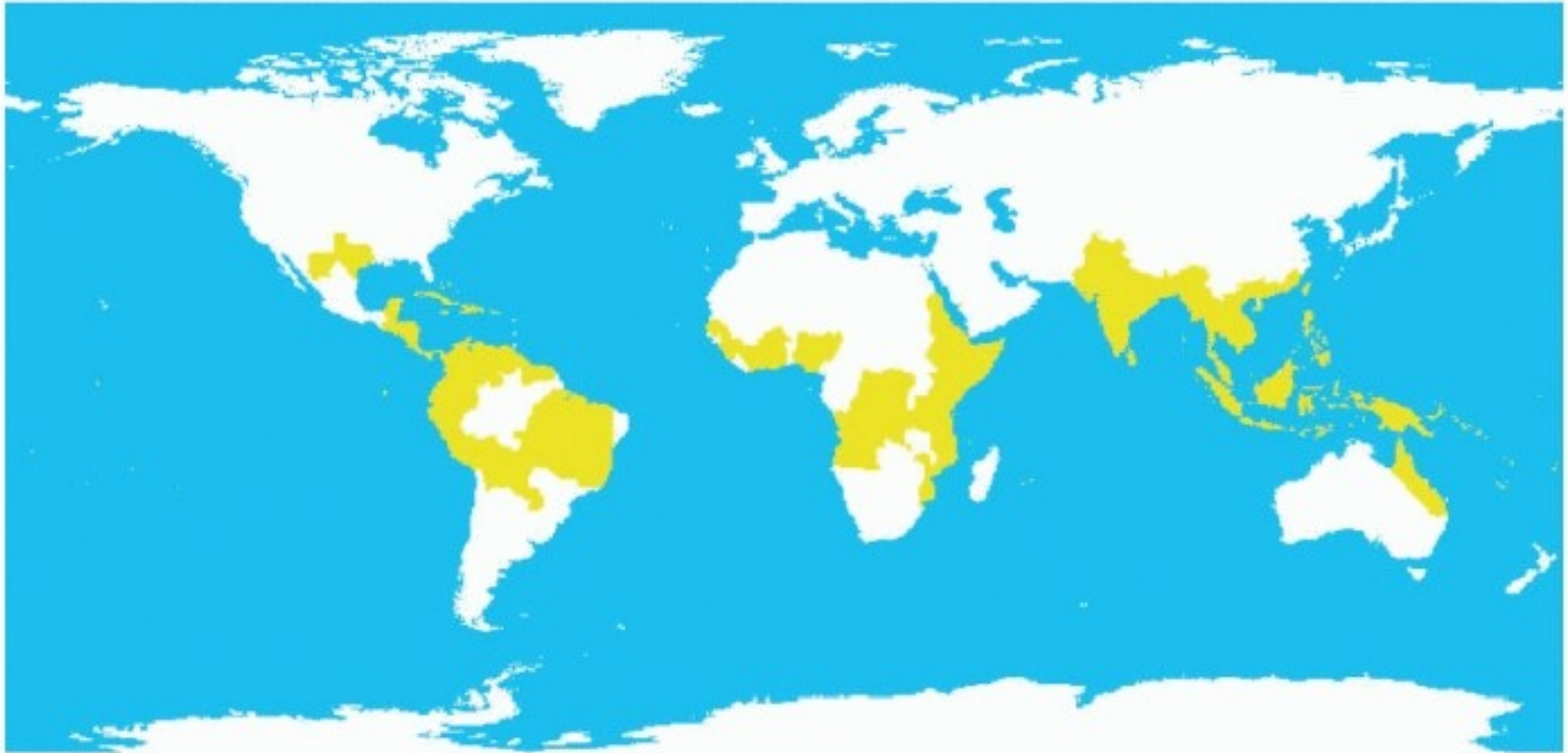
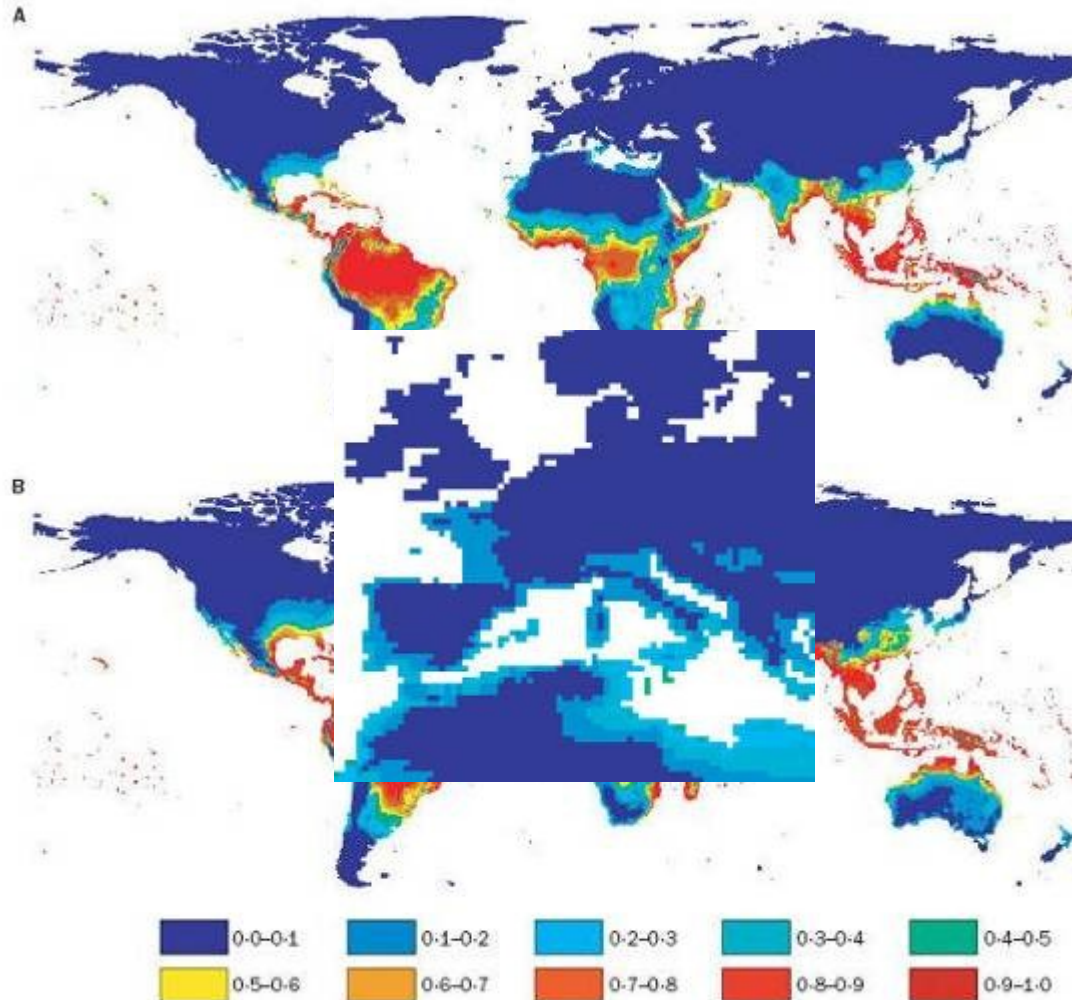


Figure 1: **The distribution of dengue fever between 1975 and 1996**  
In yellow areas there was at least one outbreak between 1975 and 1996.<sup>3</sup>

# ...Prospective 2085 ?



*Hales S, Lancet 2002*

Figure 2: Estimated baseline population at risk in 1990 (A) and estimated population at risk in 2085 (B)

Results of a logistic regression model with vapour pressure (humidity) as the predictor of dengue fever risk, using climate data from 1961 to 1990 (A). Forecast geographical distribution of dengue transmission based on climate projections for 2080–2100 from a global circulation model (CCGMA2) (B). Colours represent probability of dengue fever transmission.

# Fièvre de la vallée du Rift

## **Climate and Satellite Indicators to Forecast Rift Valley Fever Epidemics in Kenya**

Kenneth J. Linthicum,<sup>1\*</sup> Assaf Anyamba,<sup>2\*</sup> Compton J. Tucker,<sup>2</sup>  
Patrick W. Kelley,<sup>1</sup> Monica F. Myers,<sup>2</sup> Clarence J. Peters<sup>3</sup>

*Science, July 1999*



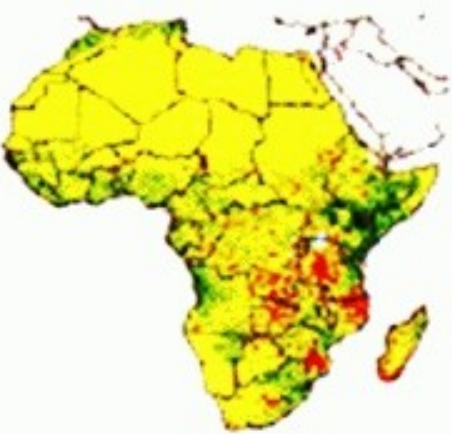
# REPORTS



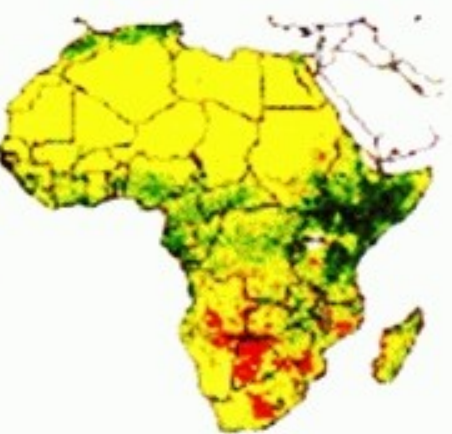
SEPTEMBER 1997



OCTOBER 1997



NOVEMBER 1997



DECEMBER 1997



JANUARY 1998



FEBRUARY 1998



MARCH 1998



APRIL 1998

## NDVI ANOMALIES



# Sentinelles

*Country Profile : France. Lancet, 1997*

1200 GPs

Volontaires, bénévoles

Transmission électronique

Protocoles standardisés

**Grippe**

**Diarrhées**

**Varicelle, Zona**

**Oreillons, Rougeole**

**Hépatite-A, Hépatite-B**

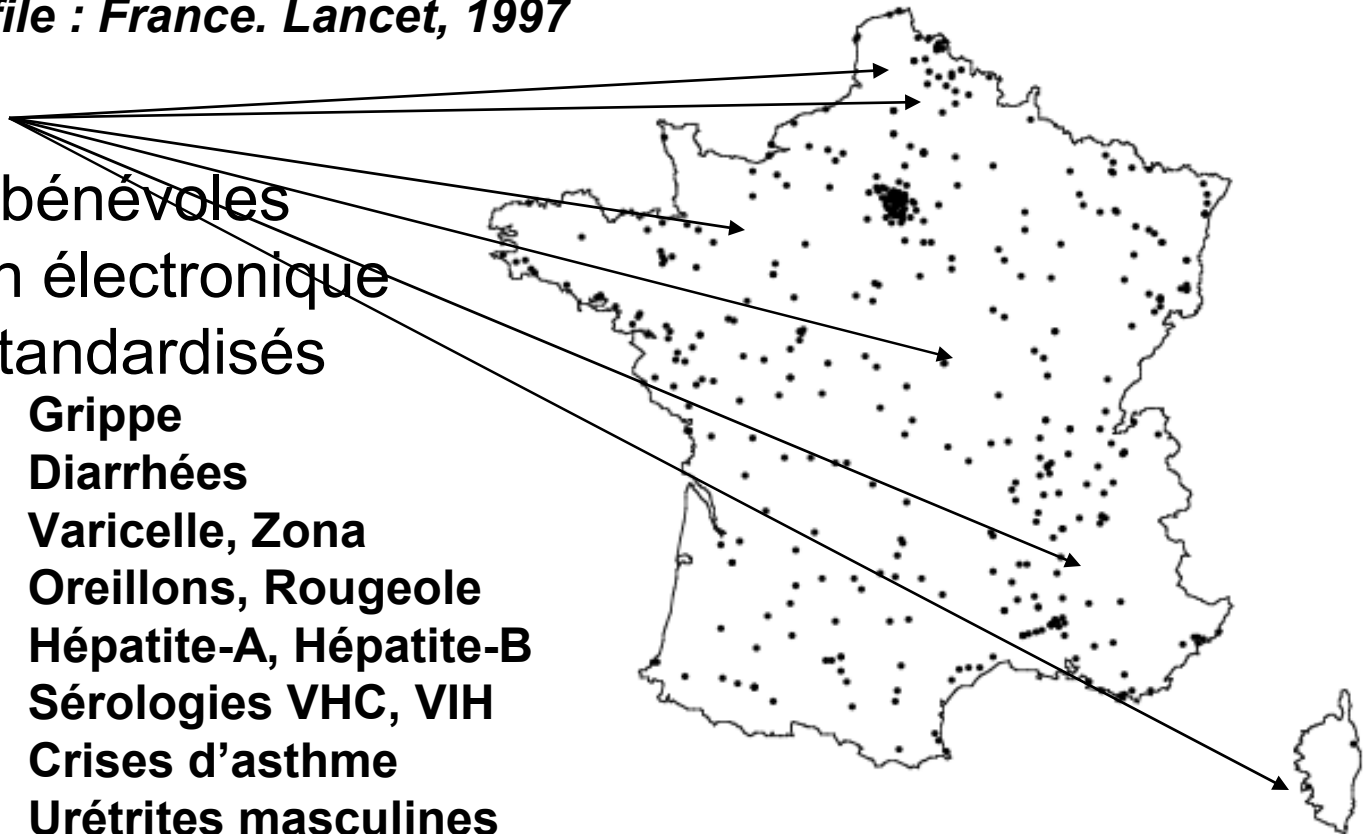
**Sérologies VHC, VIH**

**Crises d'asthme**

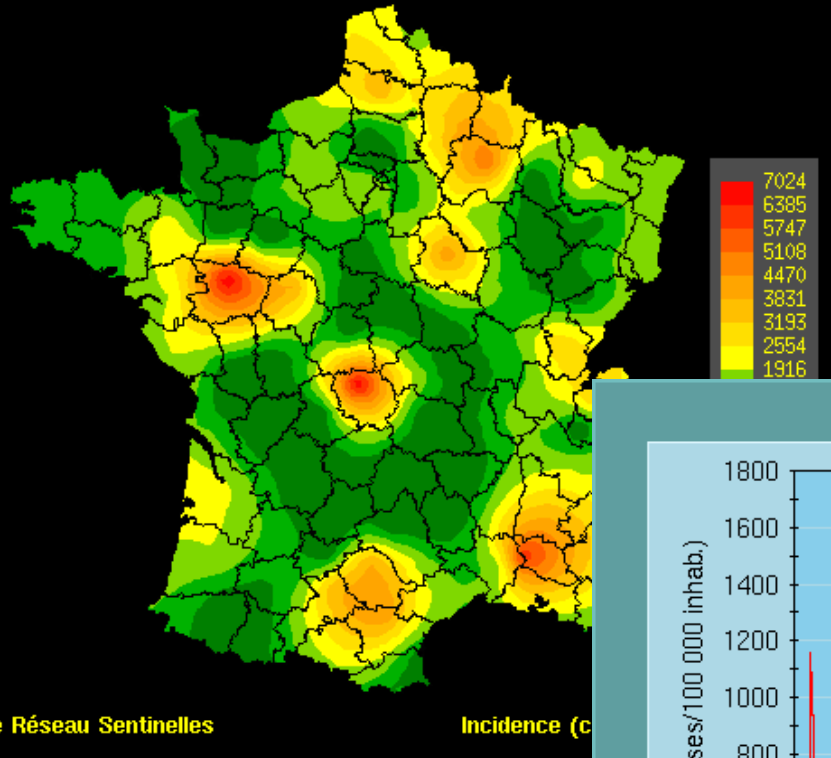
**Urétrites masculines**

**Tentatives de suicides**

**Hospitalisations**

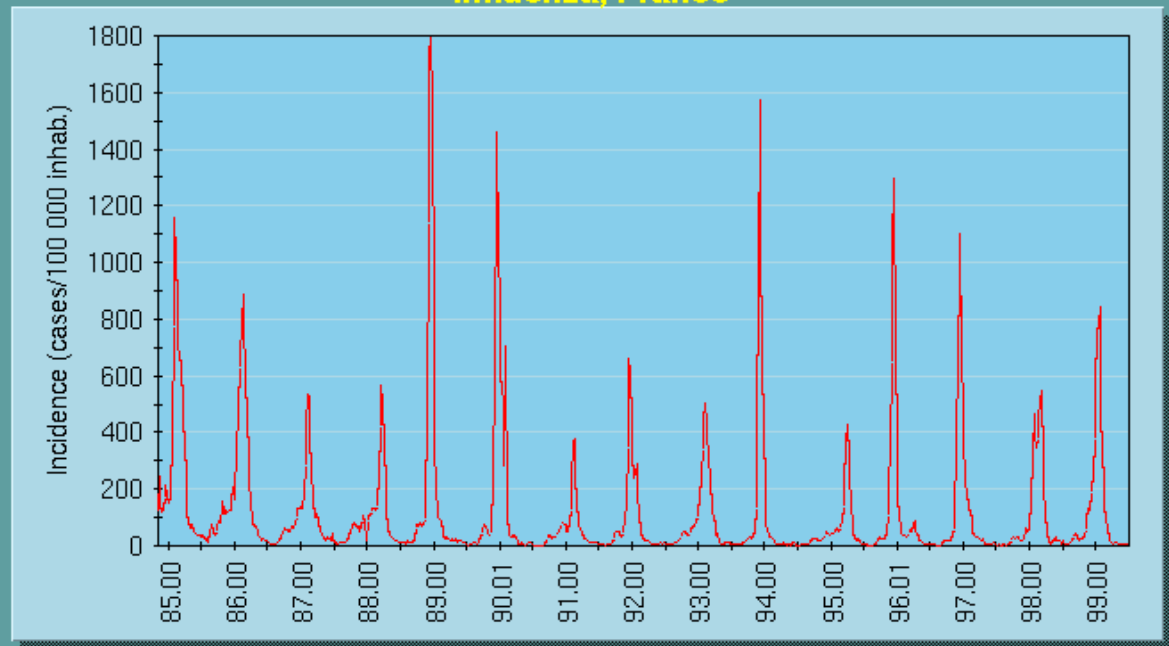


Influenza – 99 Jan



# Observation de la grippe en France

Influenza, France

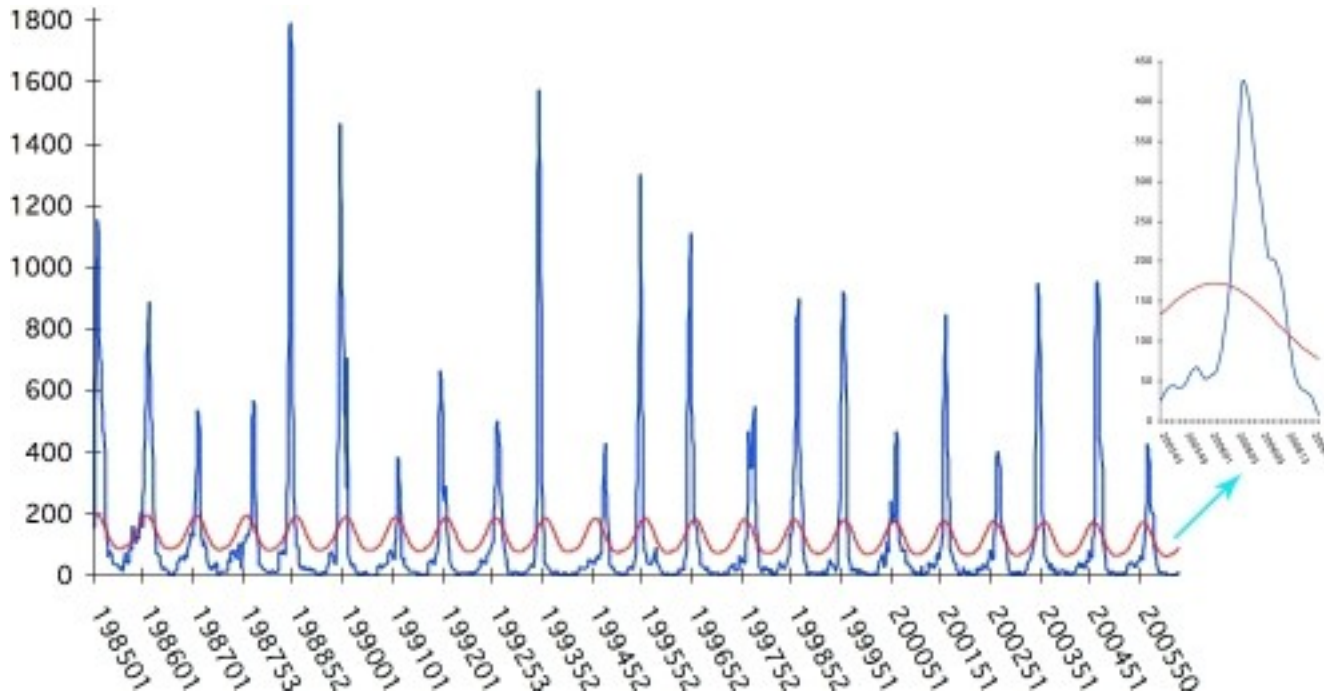


© Le Réseau Sentinelles

*Carrat, Am J Epidemiol 1991*

*Costagliola, Am J Pub Health 1991*

# Les épidémies de grippe sont prévisibles



Chaque hiver  
en moyenne :  
2,5 M de cas  
6 M de grippe  
6000 décès

# Variabilité épidémique

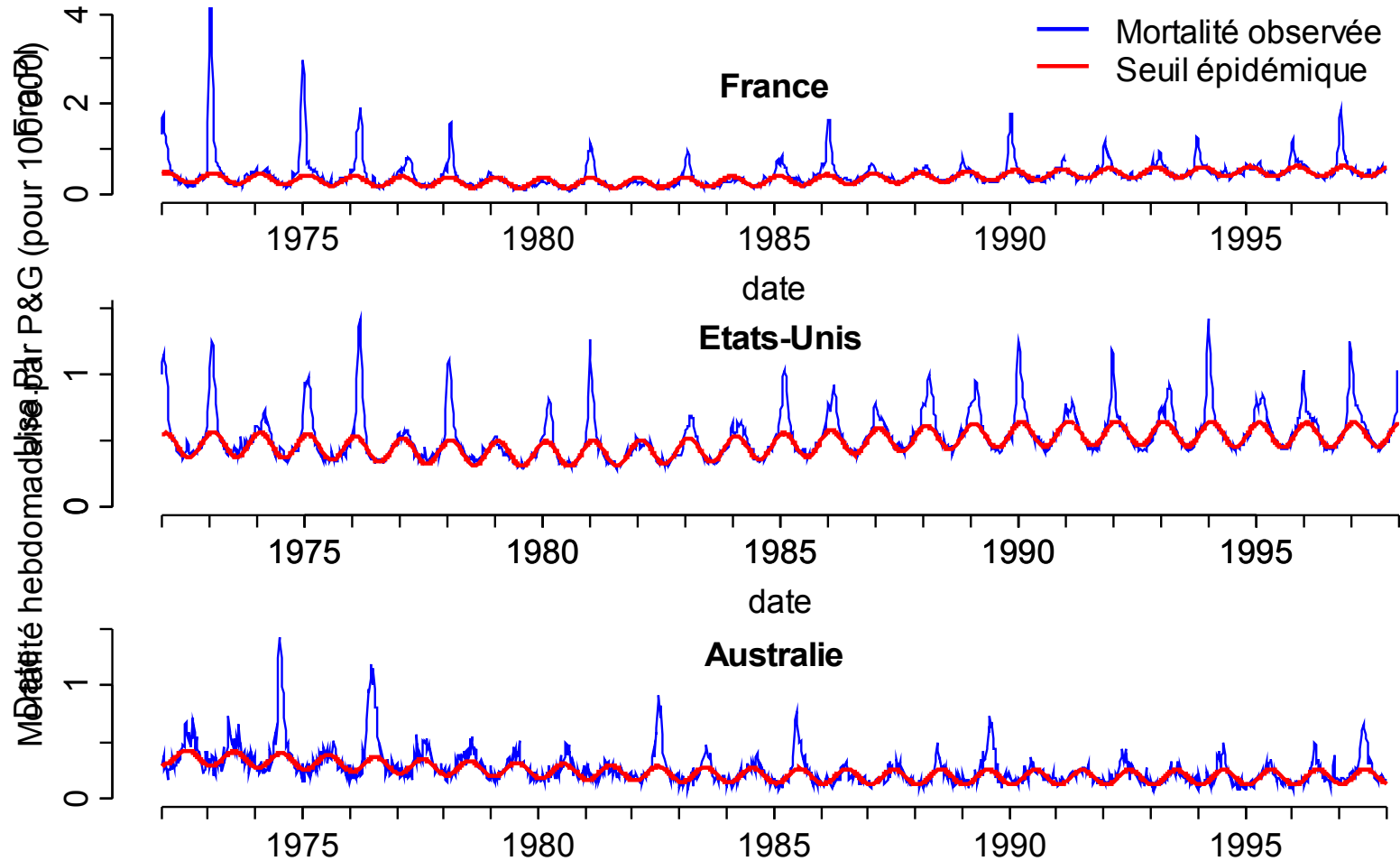
Depuis 1984 :

- Début d'épidémie : 6 nov. (2003) à 20 mars (1995)
  - Durée de l'épidémie : 4 sem. (1990) à 13 sem.(1985)
  - Taille de l'épidémie : 700 000 (1990) à 4,6 M (1989)
  - Souches virales circulantes : A(H3N2), A(H1N1), B
- => Pas de facteur déclenchant connu à ce jour.

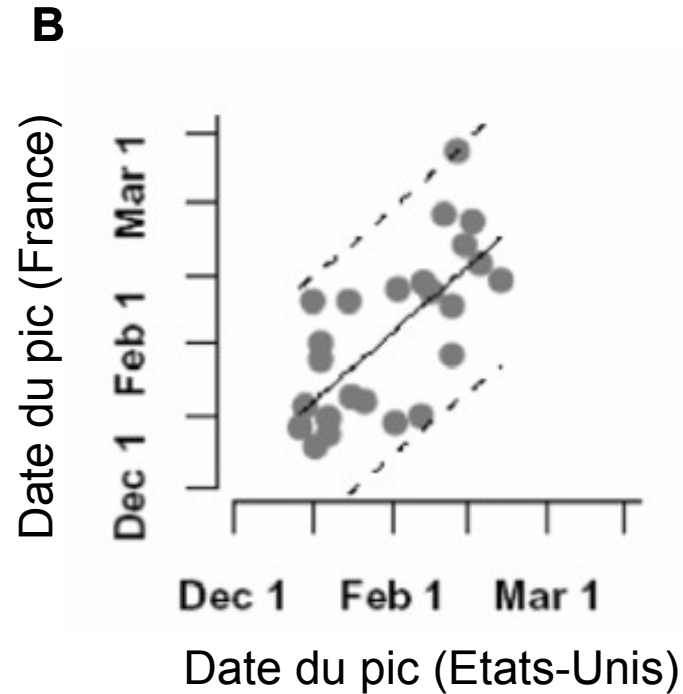
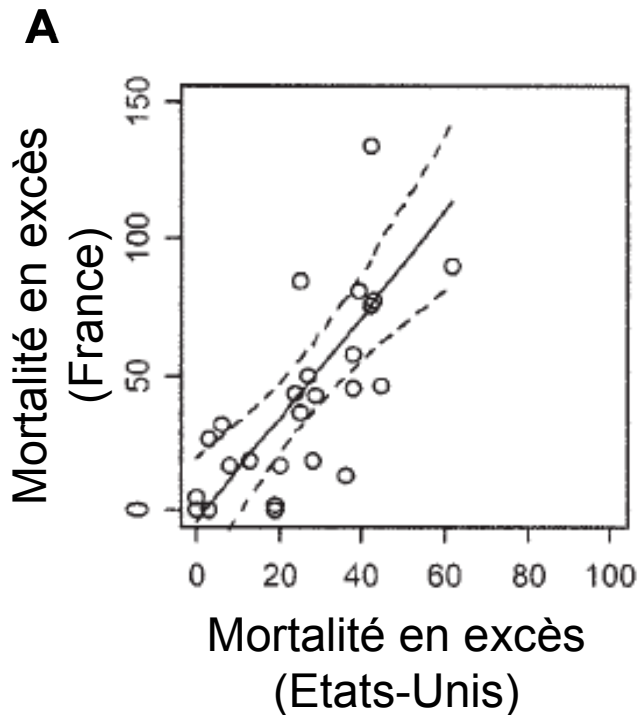
# Prévoir l'imprévisible ?

1. Quelles sera la taille de l'épidémie cet hiver ?
2. Lorsque l'épidémie aura démarré, quelle sera son évolution et ses caractéristiques ?
3. Quelle sera l'efficacité du vaccin au cours de l'épidémie ?

# Mortalité en excès par grippe et pneumonie France-USA-Australie



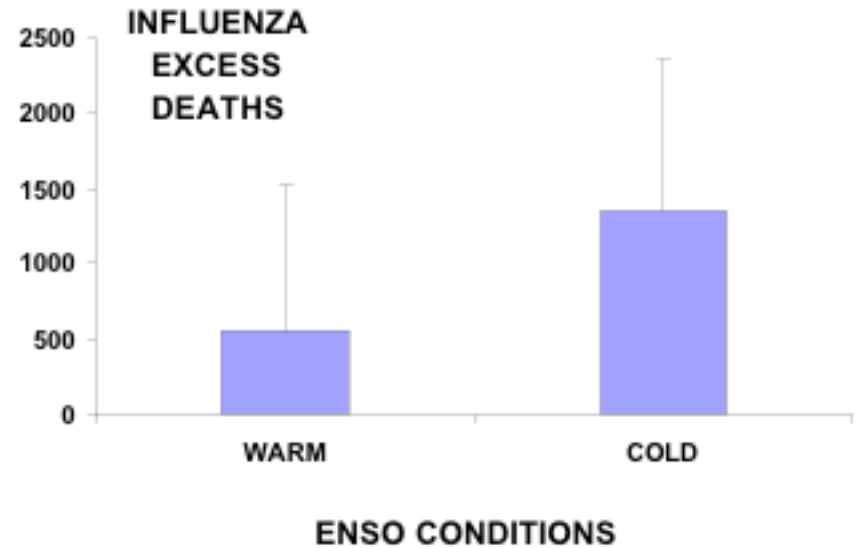
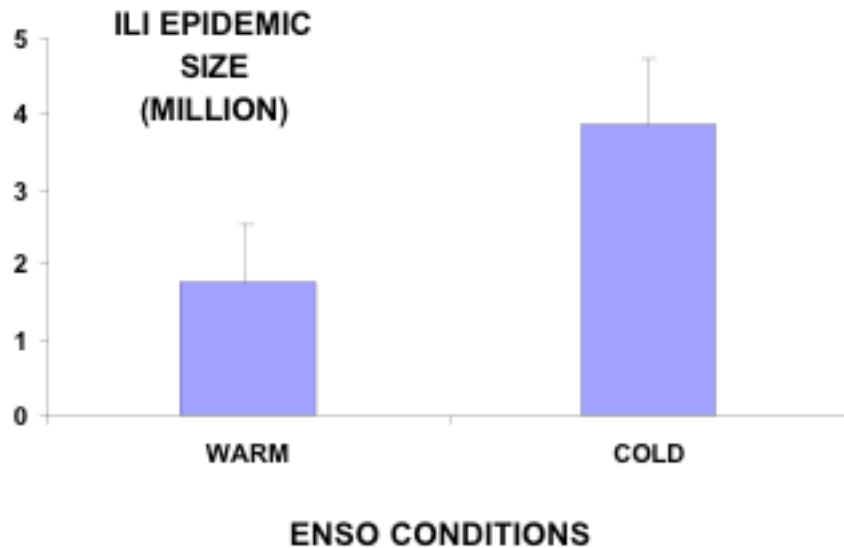
# Synchronisme des épidémies de grippe (mortalité)



Délai moyen France-USA : 0,5 semaine ( $P < 0,001$ )

*Viboud, EID, 2004*

# La Nina = fortes épidémies



*Flahault A, et coll. ICS 2004;1263:73-77.  
Viboud C et coll. Eur J Epidemiol, 2004*

# Les méthodes

## Prévoir l'épidémie avant qu'elle ne soit arrivée

Modèle de régression linéaire multiple expliquant la taille de l'épidémie (morbidity), avec les types et sous-types viraux prédominants et la valeur de l'indicateur climatique MEI comme co-variables. Le modèle a été validé sur 22 saisons grippales et prédit la taille de la 23ème.

## Prévoir l'évolution d'une épidémie en cours

Méthode des analogies : recherche dans la série chronologique passée du vecteur "plus proche voisin" du vecteur constitué des trois semaines courantes. Préviation des semaines futures à partir de l'expérience passée (moyenne des évolutions des plus proches voisins)

# Les données utilisées

- Données de morbidité (syndromes grippaux) : France 1984-2006, réseau Sentinelles (<http://www.sentiweb.fr>)
- Données de virologie (types/sous-types prédominants) : Institut Pasteur, Paris, Université de Lyon, et FluNet (OMS)
- MEI (Multivariate Enso Index)
- Couverture vaccinale par âge

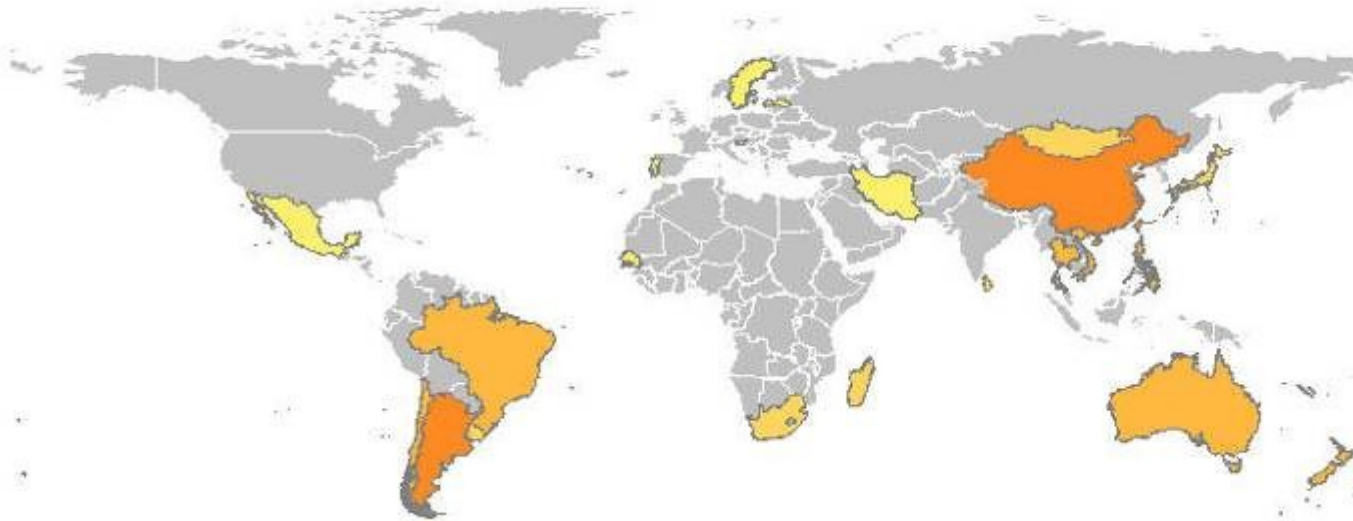
# **Deux « paris » nécessaires**

Quelle(s) souche(s) prédominera(ront)  
cet hiver ?

Quel niveau aura l'indicateur climatique  
cet hiver ?

# L'hiver austral (juillet-août 06)

Communicable Diseases --> Influenza -->  
Isolates obtained, Total, Sum over all  
periods, From week 26/2006 to week  
53/2006  
Source: FluNet

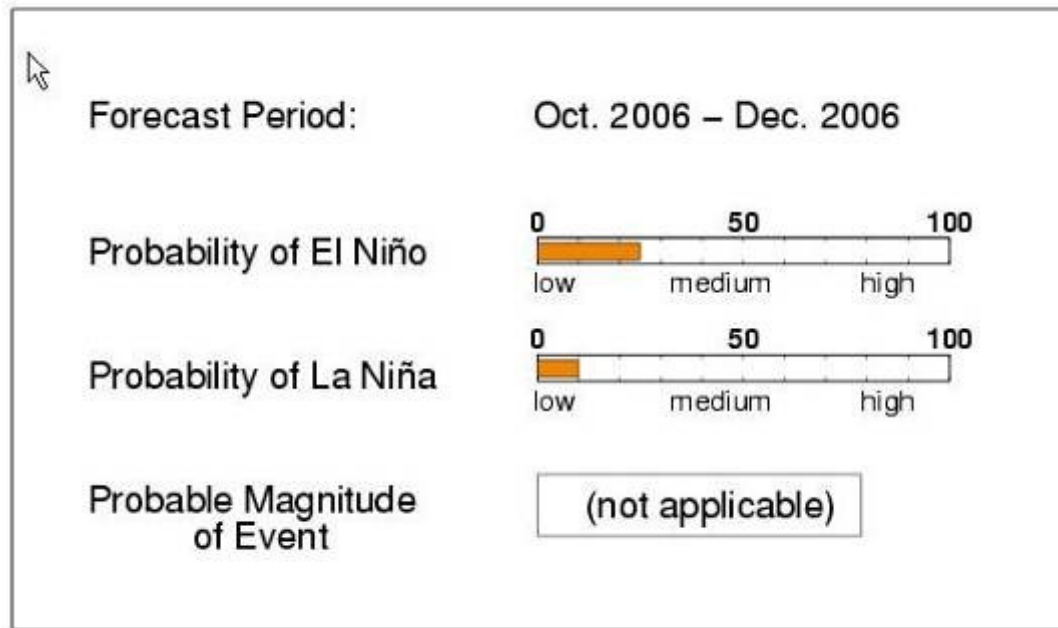


**B, H3N2**

*Source : FluNet, OMS*

# El Nino modéré (automne 06)

MEI = 2,00



[http://iri.columbia.edu/climate/ENSO/currentinfo/SST\\_table.html](http://iri.columbia.edu/climate/ENSO/currentinfo/SST_table.html)

Année	MEI cumulé	Virus dominants	Taille observée	Modèle	IC95%
2002	3,84	B,A/H3N2	1,6	1,8	[0,9;2,7]
2003	1,92	A/H3N2	2,9	3,0	[2,4;3,6]
2004	2,61	A/H3N2	3,1	2,9	[2,3;3,5]
2005	-0,95	B,A/H1N1	1,6	1,6	[0,5;2,7]
2006	2,00 ?	B,A/H3N2 (?)		2,2	[1,3;3,0]

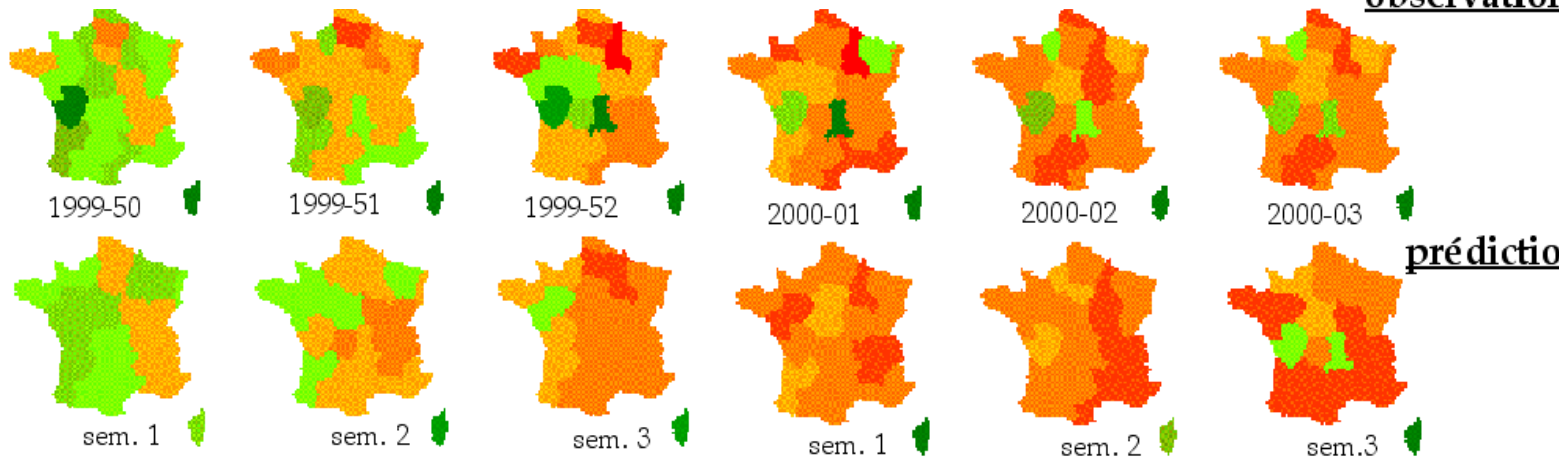
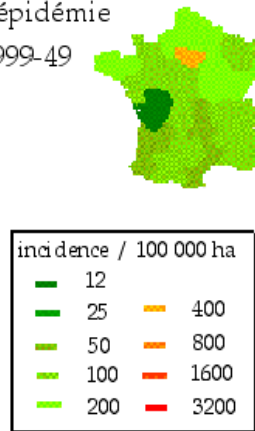
2005-2006 : 3,2 millions de cas prédits...

2006-2007 : 2,2 millions de cas ?

# Evolution d'une épidémie en cours : météo-épidémiologie

semaine de début d'épidémie  
1999-49

## La méthode des analogies



*Viboud C. et coll. Am J Epidemiol, 2003*

# Résumé : prévoir la grippe

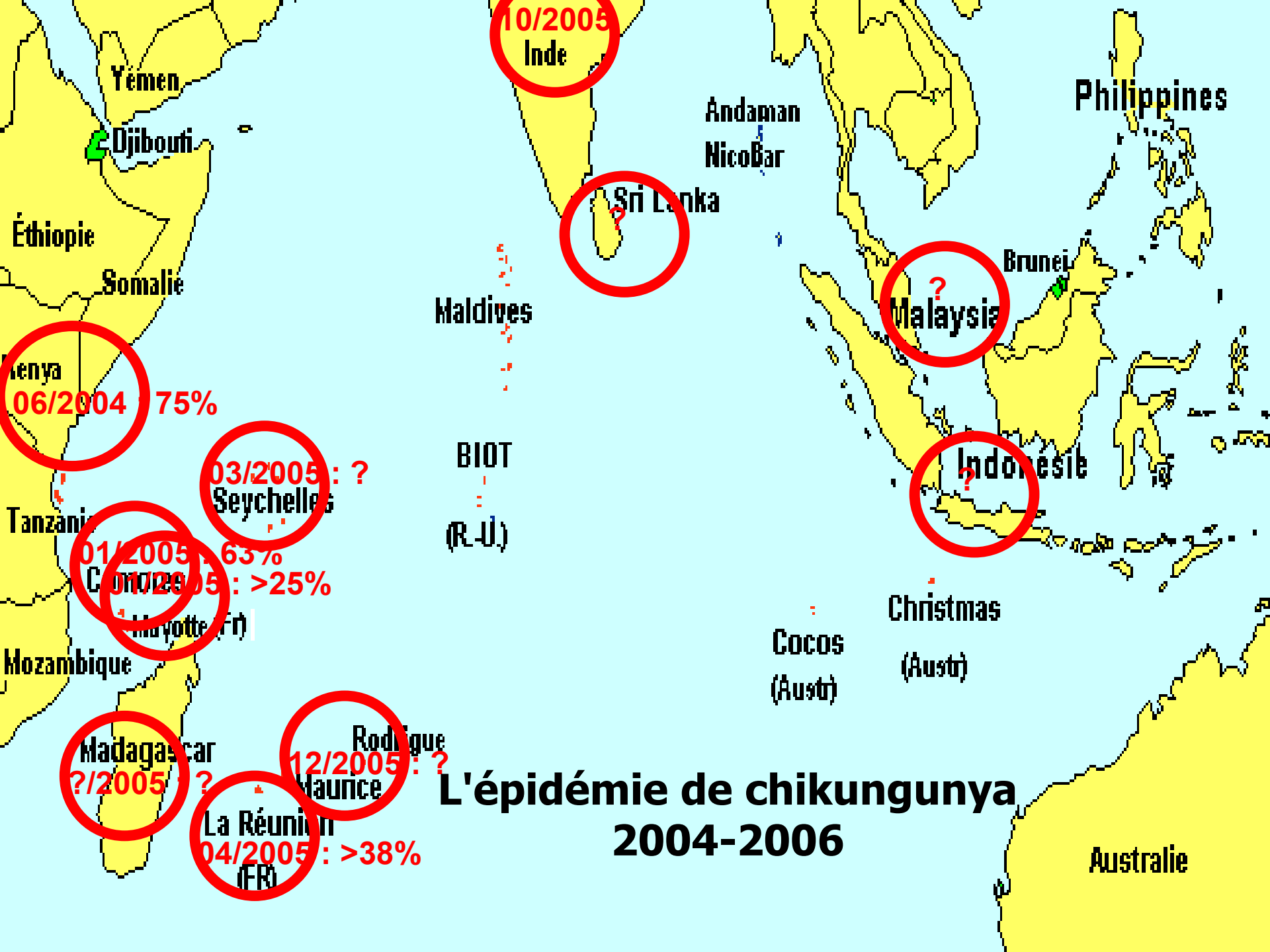
**67% de la variabilité de la taille épidémique s'explique par la (les) souche(s) virale(s) prédominante(s) et l'indicateur climatique**

=> Prédire au préalable

- la circulation virale de la saison à venir
- le niveau de l'oscillation ENSO (durant l'automne).

# ...Prospective 2100 ?

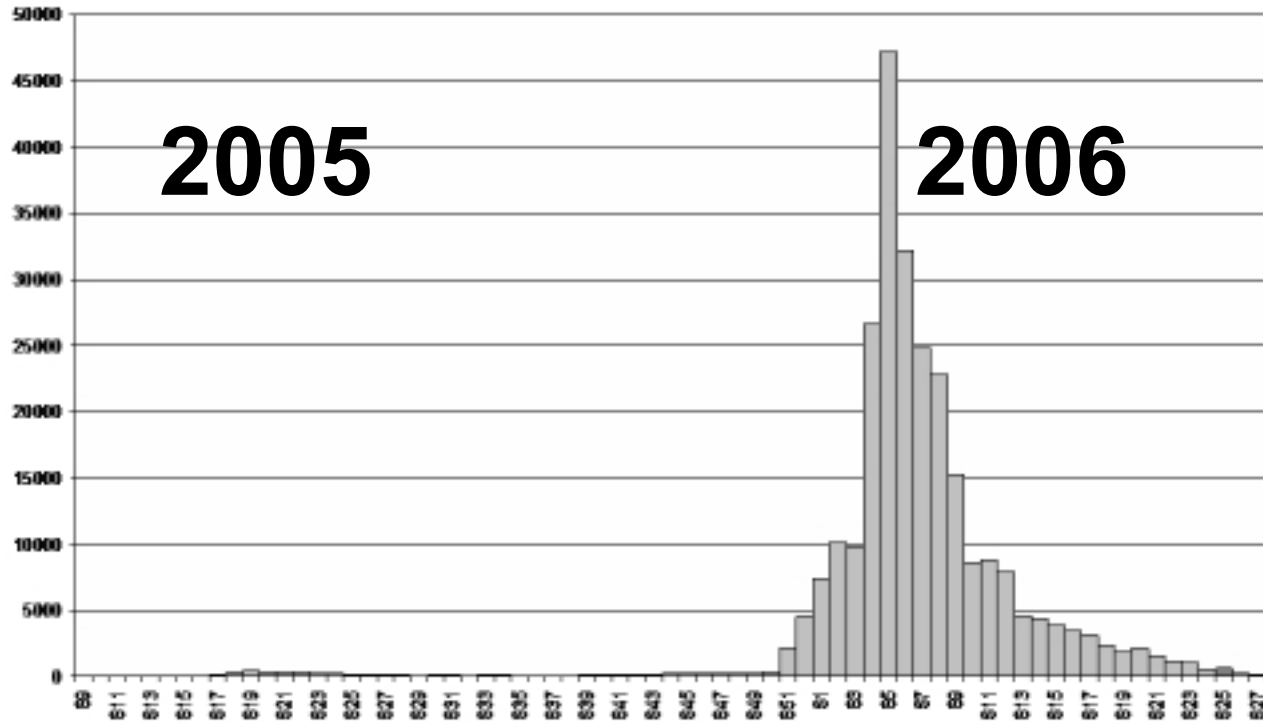
- Les dynamiques épidémiques sont complexes et ne peuvent se résumer en de simples interprétations statistiques
- Demain, il faudra incorporer les données satellitaires dans des modèles biologiques explicatifs de la transmission du pathogène
  - liaison entre mortalité et abondance du vecteur et données satellitaires
  - développements de modèles biologiques (ex. trypanosomiase, mouche tsé-tsé, tique africaine)
- Il est urgent d'abord de mieux comprendre les dynamiques de transmission



# L'épidémie de chikungunya 2004-2006

Australie

# Deux vagues épidémiques en 2005 et 2006

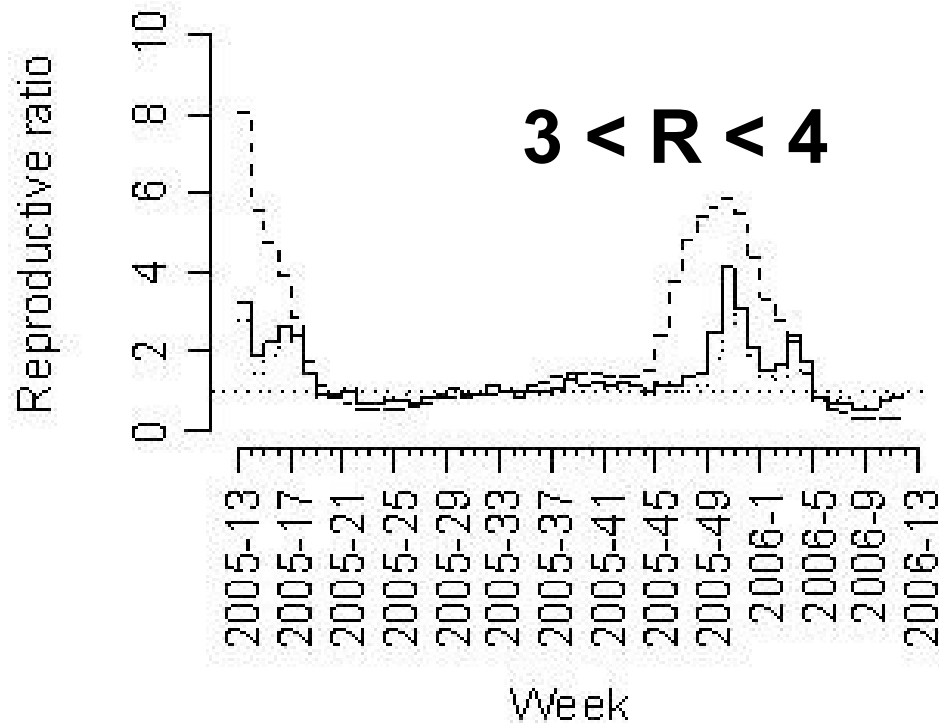
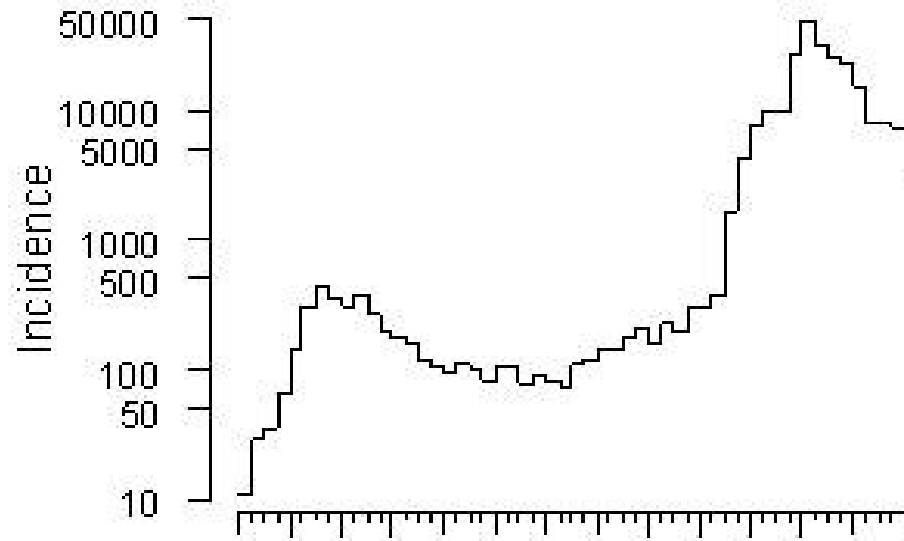
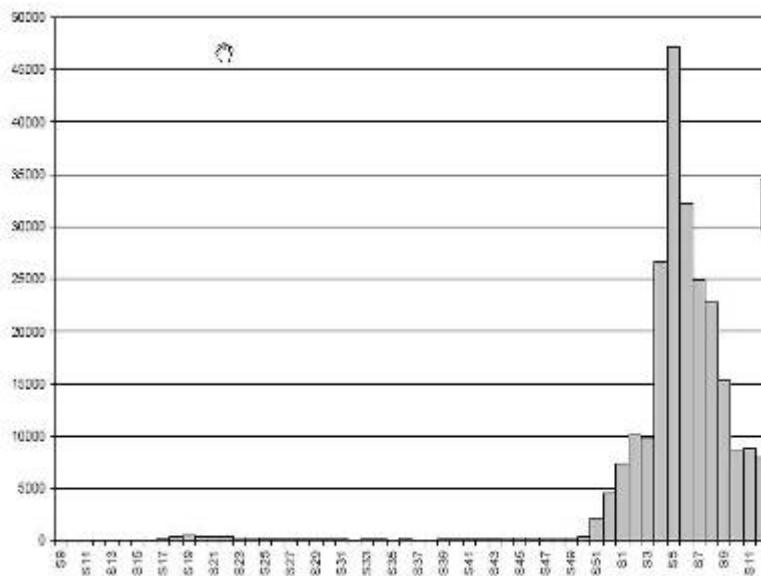


# Séquençage du génome viral du chikungunya

Océan Indien : 92 séquences de 89 patients

Date	Séquences	A226	V226
Mars-Juin 2005	19	19	0
Septembre-Décembre	27	0	27
Janvier-Mars 2006	46	6	40

Switch de A226 vers V226 entre les 2 vagues épidémiques



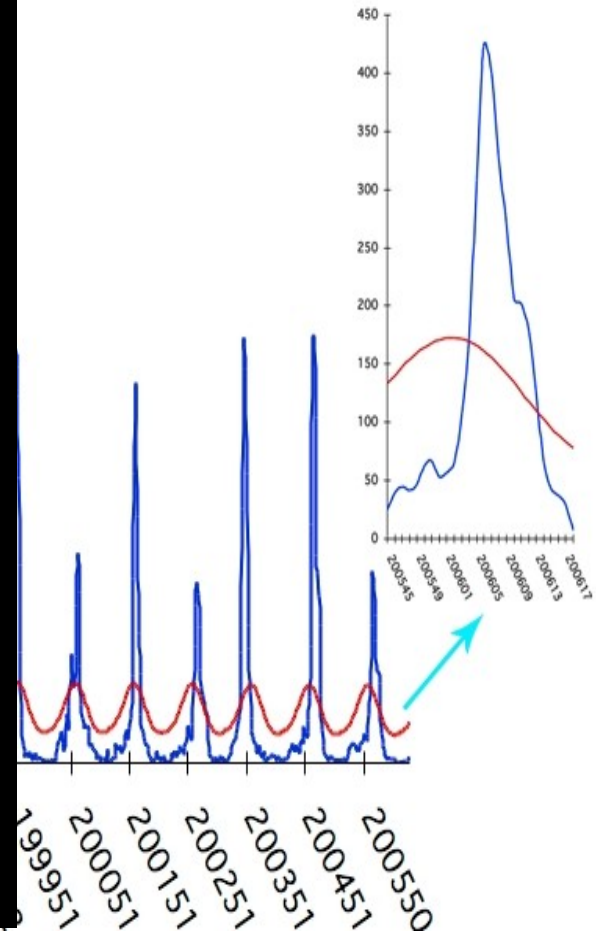
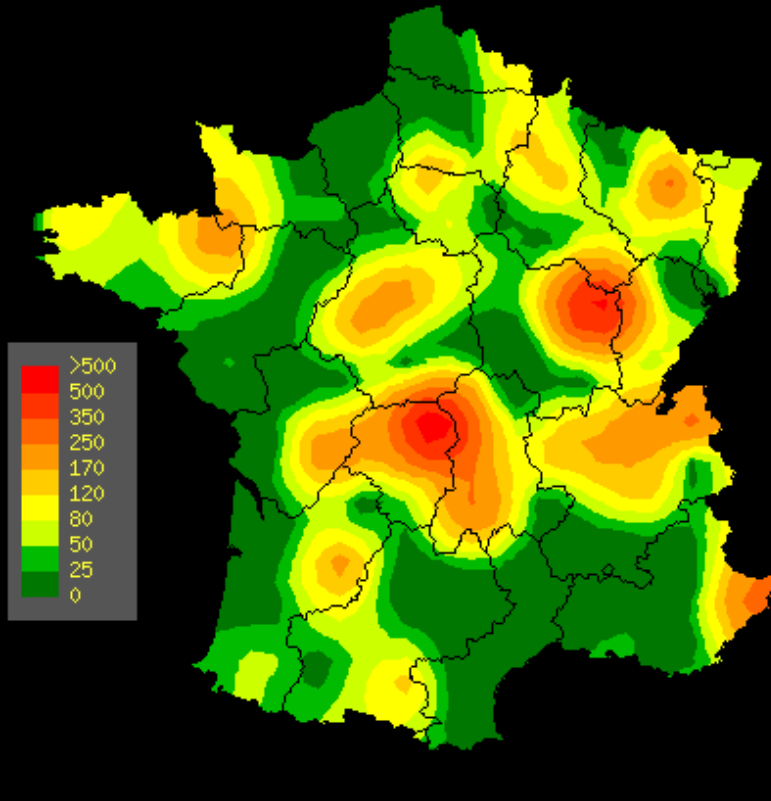
Deux vagues  
de même  
nature !

*Boelle et Flahault, 2006*

# Conclusion : la pluridisciplinarité !

- Mathématiques : la modélisation pour « évaluer » a priori des stratégies d'intervention
- Informatique : la surveillance électronique temps-réel pour anticiper, la simulation individus-centrés
- Sciences humaines et sociales
  - Réseaux sociaux, perception du risque, systèmes d'information géographique, économie
- Epidémiologie clinique : antiviraux, masques, vaccins, fermeture d'école => plus d'expérimentation
- Santé animale, entomologie, ornithologie, environnement, télédétection, microbiologie...

Grippe Semaine 2006-03  
en nombre de cas pour 100 000 habitants



Sentinelles (R)

INSERM

(c) Mapoflash

Carte d'interpolation spatiale des données

L'interpolation se fait sur la base des incidences départementales