

# Les biostatistiques : quels problèmes, quels outils ?

M. Bailly-Bechet

Université Claude Bernard Lyon I – France

Document disponible à :  
<http://pbil.univ-lyon1.fr/members/mbailly>

# Les faits : une production massive de données biologiques

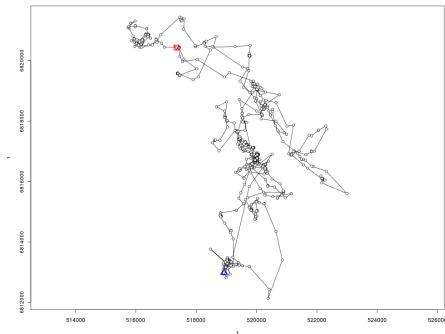
Les dernières décennies de recherche en biologie, couplées à des développements technologiques, ont permis l'étude fine de très grand nombres de sujets, par exemple :

## Quelques exemples : génomique



La séquence et le niveau d'expression de tous les gènes d'un génome dans des conditions contrôlées.

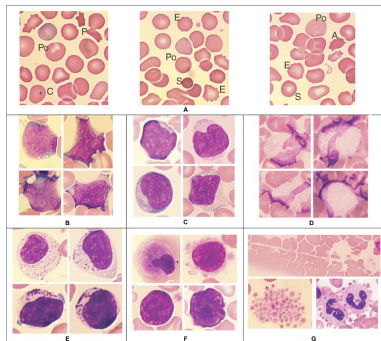
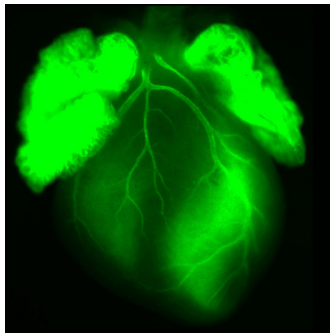
## Quelques exemples : écologie



### *Scandinavian Bear Research Project*

Les itinéraires de populations complètes d'animaux sur plusieurs années

## Quelques exemples : santé et imagerie médicale



Les études cliniques à grande échelle sur certaines affections et maladies

## Le but des biostatistiques

- ▶ **Décrire** : ordonner les données, les classer, trouver comment les décrire de manière simple.
- ▶ **Comprendre** : modéliser, proposer et tester des hypothèses explicatives, les comparer entre elles, et parvenir, grâce à certaines hypothèses, d'expliquer tout ou partie des données.

## Les problèmes particuliers posés par les biostatistiques

- ▶ Les données biologiques sont très variables, même dans des conditions contrôlées : variabilité intrinsèque du système, variabilité expérimentale, bruit dans les données. . .
- ▶ Données manquantes
- ▶ Peu de répétitions expérimentales (coûts souvent prohibitifs)
- ▶ Décisions parfois critiques dans certains domaines (notamment le médical) => problèmes associés à la valeur morale de la notion de "risque".

## Actuellement

- ▶ Développement d'une culture interdisciplinaire entre statisticiens, informaticiens et biologistes
- ▶ Cours universitaires dédiés
- ▶ Reconnaissance des besoins dans le domaine aussi bien par les universitaires que par les industries