

# Les différentes situations de multiplicité des tests Les méthodes de contrôle du risque alpha global

Pr Nicolas GIRERD  
Centre d'Investigations Cliniques Plurithématiques  
& Département de cardiologie  
Institut Lorrain du Cœur et des Vaisseaux  
CHRU Nancy

# Mise en situation



Probabilité de faire 6 (au moins une fois)  
 $\Rightarrow 1/6 = 0,17$



Probabilité de faire 6 (au moins une fois)  
 $\Rightarrow 1-(1-1/6)^2 = 1-25/36 = 11/36 = 0,31$



Probabilité de faire 6 (au moins une fois)  
 $\Rightarrow 1-(1-1/6)^3 = 1-125/216 = 91/216 = 0,42$

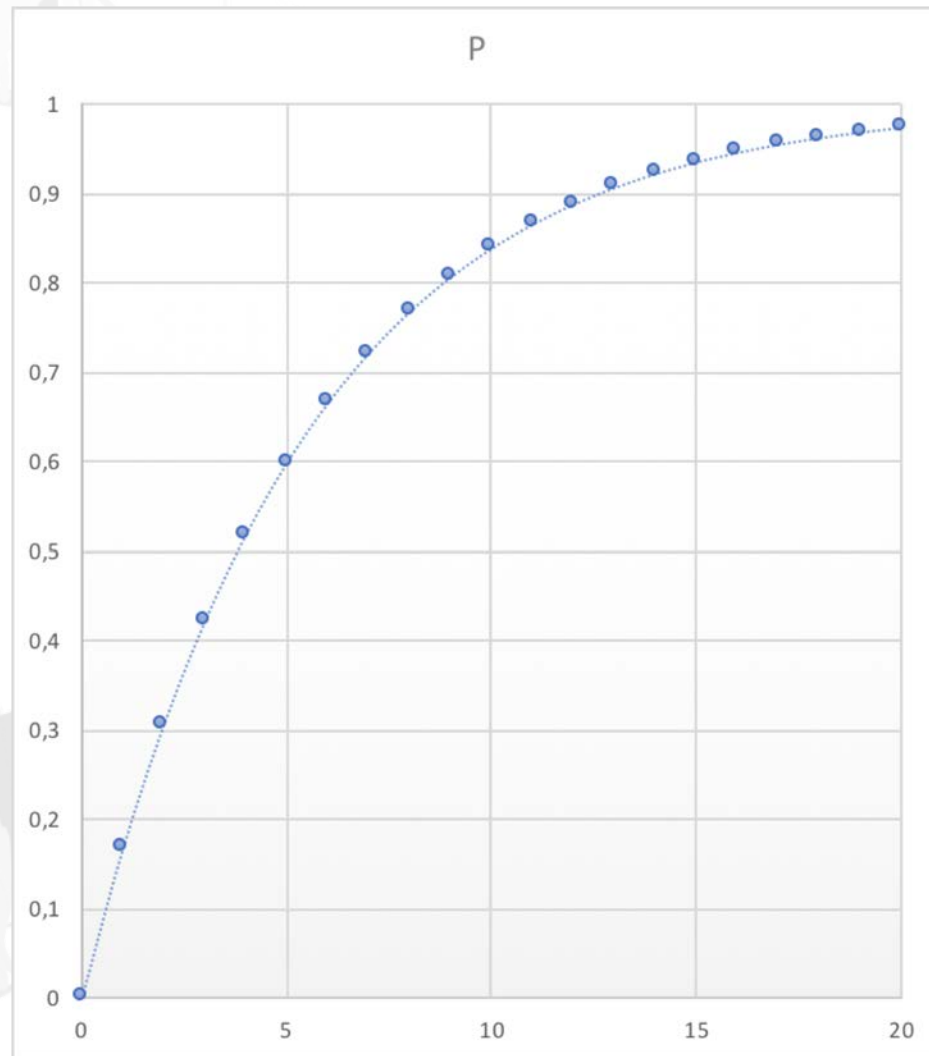
# Mise en situation

**Probabilité de faire 6  
(au moins une fois) :**

$$1 - (1 - 1/6)^n$$

**où n est le nombre de tirages**

**Tend vers 1  
>0,95 dès 17 tirages**





**100 % des gagnants  
ont tenté leur chance**

**Plus vous jouez, plus vous avez de chance de gagner !**

# Comparaisons multiples Inflation du risque alpha Familywise Error Rate

**Seuil (habituellement) de 5% pour considérer qu'un test est significatif**

On prend le risque de considérer  
que le résultat observé n'est **pas le fruit du hasard**  
lorsque la probabilité d'observer un résultat similaire ou plus éloigné de la ligne  
d'identité **du fait du hasard est inférieur à 5%.....**

**EN JOUANT UNE SEULE FOIS  
S'APPLIQUE A UN (SEUL) TEST !**

**QUELLES CONSEQUENCES DE L'UTILISATION  
D'UNE REGLE S'APPLIQUANT A UN TEST  
LORSQUE PLUSIEURS TESTS SONT UTILISES ???**

# Comparaisons multiples

## Inflation du risque alpha

### Familywise Error Rate

#### Familywise error rate (FWE or FWER)

- Probabilité d'observer au moins une conclusion erronée en réalisant plusieurs tests statistiques
- Probabilité de faire une erreur de type I
- « Family » pour famille/série de test
- Aussi appelée **erreur de type I cumulative**

**Estimation de FWER :**

$$\text{FWER} \leq 1 - (1 - \alpha_{IT})^c$$

$\alpha_{IT}$  = alpha d'un test individuel

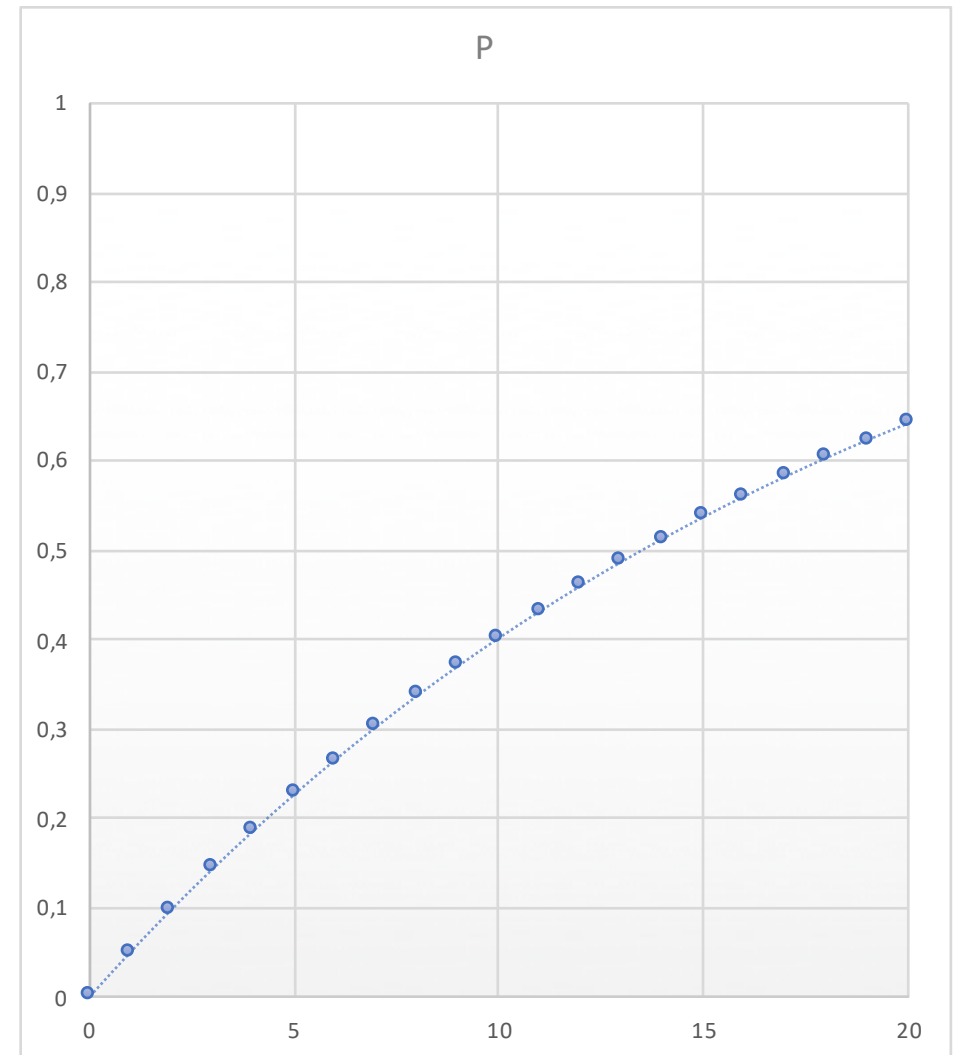
c = nombre de comparaison

# Comparaisons multiples Inflation du risque alpha Familywise Error Rate

**Probabilité de faire au moins une  
erreur de type 1 (hypothèse = 5%) :**

**Tend vers 1 avec l'augmentation du  
nombre de tests**

**9,75% à 2 tests  
22,6% à 5 tests  
40,1% à 10 tests**



# Champs d'application

## Essais cliniques

- Nombres d'outcomes considérés
- Nombres d'analyses intermédiaires considérées

## Omiques

- 100 à 5000 protéines...
- 10 000 à 4 000 000 SNPs...

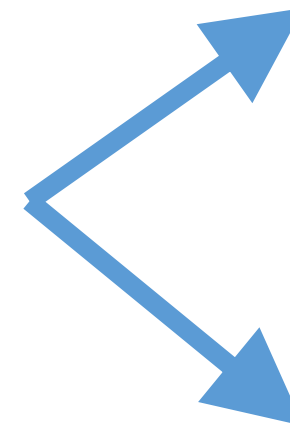
Mais surtout, la **pratique épidémiologique routinière**

# Inflation du risque alpha en pratique épidémiologique routinière

- Vous avez une base de données
    - 100 variables d'exposition
    - 5 outcomes d'intérêt
    - $100 \times 5 = 500$  tests possibles !
    - La variable 98 est très associée au critère de jugement.
- => Ecrivons (urgemment) un article à cette association !!!!



Manger de la  
confiture tue



BLACK MIRROR

# Comparaison multiple Inflation du risque alpha Familywise Error Rate

**Probabilité de faire au moins une  
erreur de type 1 (hypothèse = 5%) :**

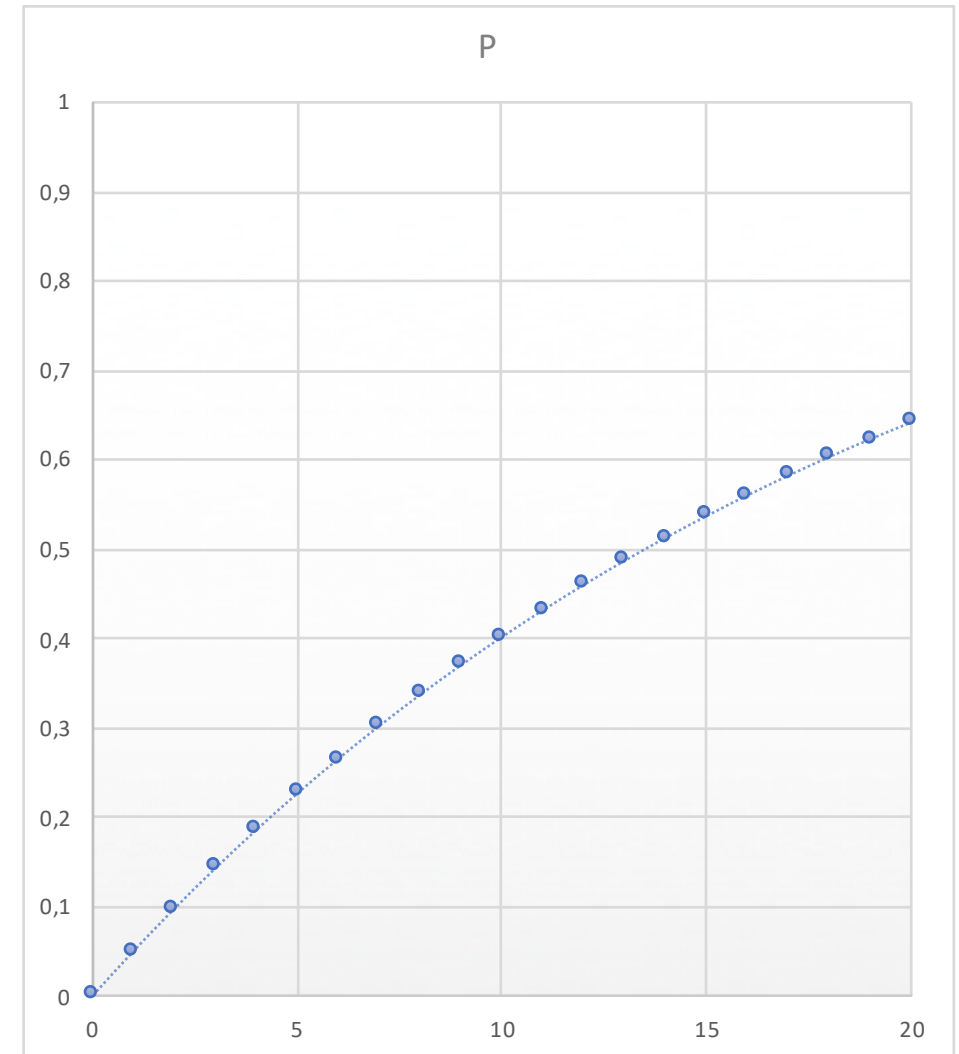
**Tend vers 1 avec l'augmentation du  
nombre de tests**

**9,75% à 2 tests**

**22,6% à 5 tests**

**40,1% à 10 tests**

**99,999 999 999 2725% à 500 tests**





# Philosophie des solutions en pratique épidémiologique routinière

# Stratégie d'analyse de cohorte

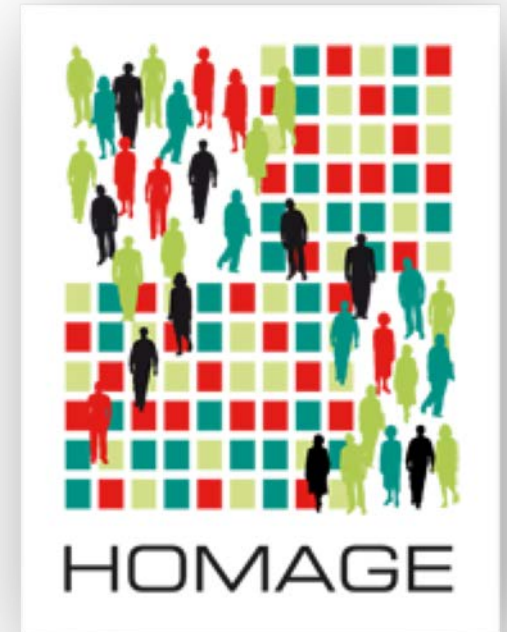
- **Adopter un seuil plus faible ?**
  - Alpha 5% =>  $\text{FWER} = 1 - (1 - 0,01)^{10} = 40,1\%$
  - Alpha 1% =>  $\text{FWER} = 1 - (1 - 0,01)^{10} = 9,6\%$
  - Mais  $c=100$  =>  $\text{FWER} = 1 - (1 - 0,01)^{100} = 63,4\%$
- **Avoir une hypothèse !!!**
- Tests seulement les hypothèses **préspécifiées**
- Planifier/limiter les analyses afin de **limiter la multiplicité des tests**
  - **SAP de cohorte à publier ?**
  - Les analyses non répertoriées dans ce SAP princeps sont exploratoires
- Etre **PRUDENT** dans l'interprétation des résultats

# Stratégie d'analyse de cohorte

- Chercher une validation interne
  - Associations avec plusieurs outcomes explorant le même concept
  - Ex : A est associé à une variable de dysfonction diastolique mais à aucune autre....  
Très différent de A est associé à toutes les variables de dysfonction diastolique.
- Chercher une validation externe
  - i.e. dans une autre cohorte !
  - **Non les reviewers/editeurs ne cherchent pas (que) à vous descendre, la validation externe est la base de l'avancée des connaissances en médecine !**

# Inflation du risque alpha dans le champ des « omiques »

- **Cohorte STANISLAS**
  - 4200 sujets suivi 20 ans
  - Lors de la dernière visite
    - Evaluation clinique = > 100 variables
    - Evaluation vasculaire = >20 variables
    - Echocardiographie de recherche = >50 variables
    - Protéomique = 500 variables
    - GWAS = 526,785 SNPs
  - **Probabilité d'obtenir une association significative sans correction >99%**





# Philosophie des solutions dans le cadre des analyses omiques

# Stratégies d'analyse « omiques »

- **Historiquement, la méthode/correction de Bonferroni**
  - Bonferroni, mathématicien italien (1892-1960)
  - Adaptation du seuil considéré comme significatif
    - **0,05/c** ou c est le nombre de comparaison
    - $\text{FWER} \leq 1 - (1 - \alpha/c)^c$
    - Exemple pour 10 tests => Seuil à 0,005 /  $1 - (1-0,005)^{10}=0.049$
    - Exemple pour 100 tests => Seuil à 0,0005 /  $1 - (1-0,0005)^{100}=0.049$
    - **Problèmes résiduels**
      - Hypothèse de tests indépendants
      - Seuils très conservateurs... Perte de puissance => Inflation du risque Beta !
  - Les « corrections » de la correction
    - Holm-Bonferroni, Sidak-Bonferroni.

# Stratégies d'analyse « omiques »

- **Plus actuel, le FDR i.e. « False Discovery Rate »**
  - Formalisé par Yoav Benjamini et Yosef Hochberg en 1995
    - La méthode de Benjamini Hochberg (BH) est la plus utilisée actuellement
  - Conceptuellement différent
    - **Contrôler la proportion de découvertes (rejetant  $H_0$ ) attendues qui sont fausses (faux positif).**
  - Méthode **moins drastique** que les méthodes de control des familywise error rate....  
**Au prix d'une augmentation des erreurs de type I...**

# Stratégies d'analyse « omiques »

- **Plus actuel, le FDR i.e. « False Discovery Rate »**
- Procédure de BH ascendante linéaire
  - Pour un nombre  $m$  de tests,
  - Ordonner les p-values « classiques »  $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(m)}$
  - Pour un FDR au niveau  $q$  donné, déterminer

$$k = \max \{i : p_{(i)} \leq (i / m)q\}$$

- On rejette les hypothèses nulles :  $H_{(1)}^0, H_{(2)}^0, \dots, H_{(k)}^0$

# Stratégies d'analyse « omiques »

- Exemple de procédure de BH ascendante linéaire

- Ordonner les p-values puis déterminer  $k = \max\{i: p_{(i)} \leq (i/m)q\}$

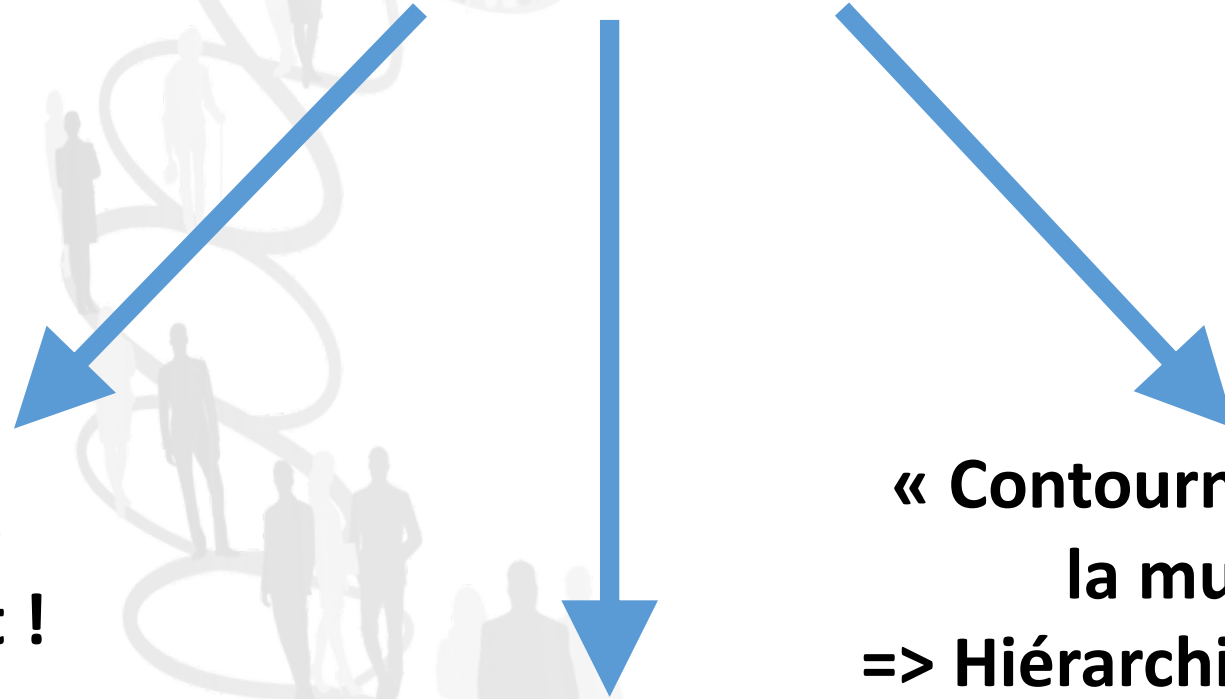
i	P value « classique »	(i/m)q	H0
1	0,0001	1/10*0,05 = 0,005	rejet
2	0,002	2/10*0,05 = 0,01	rejet
3	0,01	3/10*0,05 = 0,015	rejet
4	0,015	4/10*0,05 = 0,02	rejet
5	0,03	5/10*0,05 = 0,025	accepte
6	0,05	6/10*0,05 = 0,03	accepte
7	0,08	7/10*0,05 = 0,035	accepte
8	0,10	8/10*0,05 = 0,04	accepte
9	0,12	9/10*0,05 = 0,045	accepte
10	0,20	10/10*0,05 = 0,05	accepte

**K=4**



# Philosophie des solutions dans le cadre d'un essai clinique

## Problématique de l'inflation du risque alpha dans les essais cliniques



**Eviter la multiplicité  
=> Ne faire qu'un test !**

**« Contourner/ordonner »  
la multiplicité  
=> Hiérarchisation des tests**

**Prendre en compte  
la multiplicité  
=> Répartition du risque alpha**

Merci !

