

Taille d'échantillon dans les études en « cluster » (grappe)

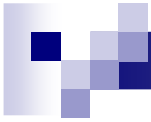
Dagmar Haller-Hester
Chef de clinique scientifique
Service de médecine de premier
recours



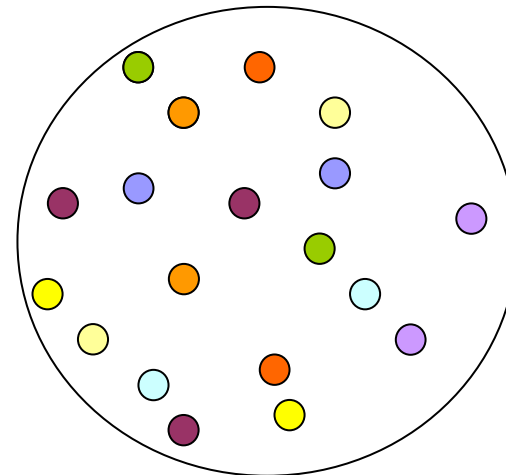
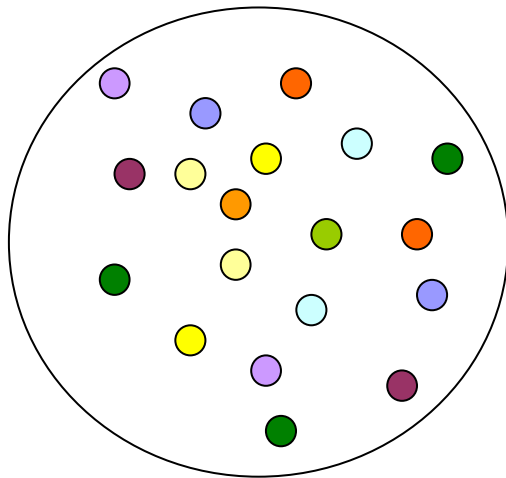
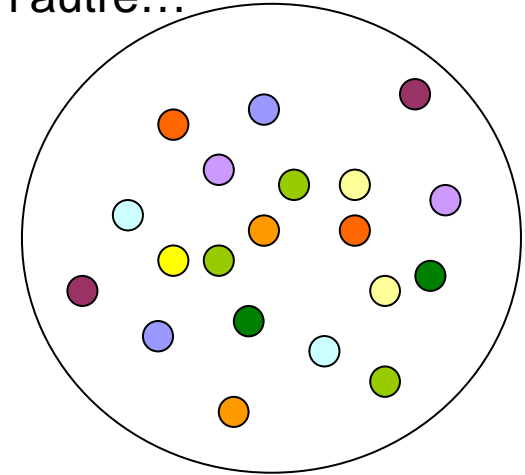
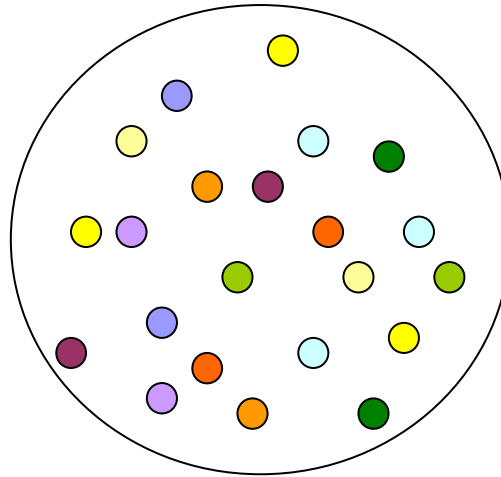
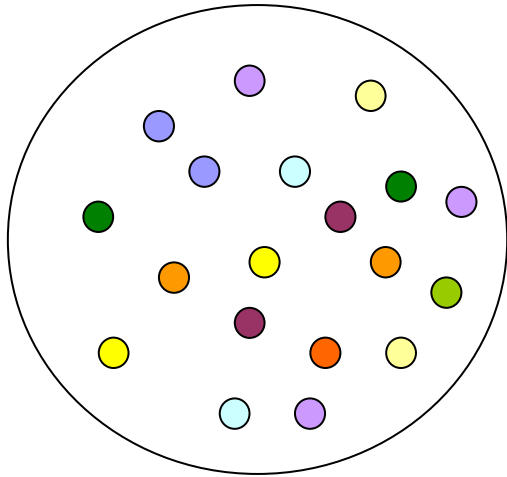
Dessins d'études en « cluster »

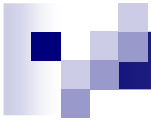
- Essais cliniques ou études observationnelles
- L'échantillonnage primaire ne se fait pas au niveau individuel mais au niveau de groupes (« en grappe »)



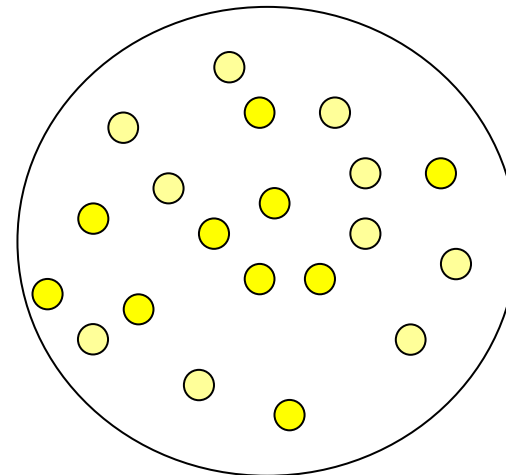
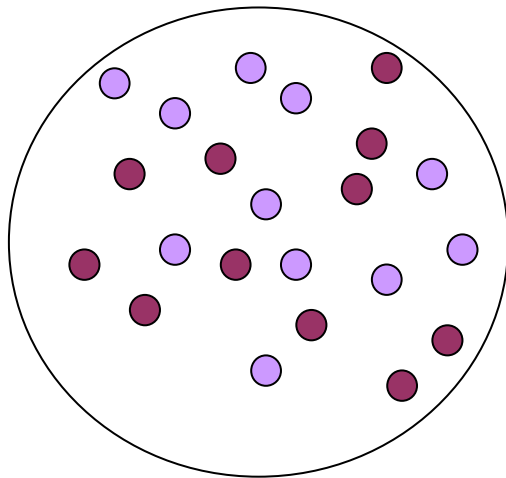
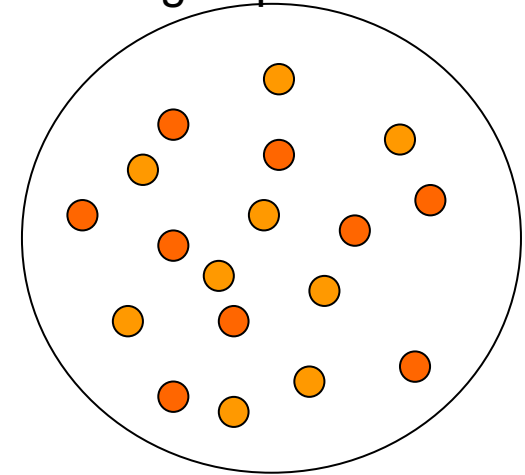
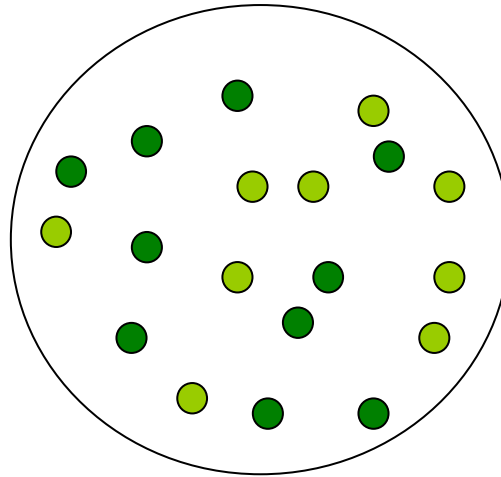
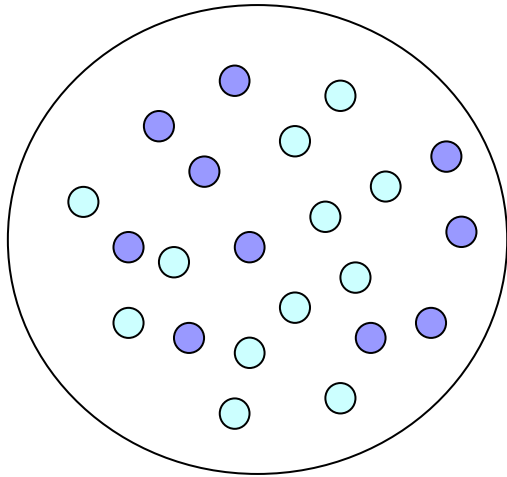


On pense souvent que la diversité des individus est identique d'un groupe à l'autre...





En réalité, les individus au sein d'un même groupe se ressemblent plus que s'ils viennent de groupes différents





Exemple (1): les individus peuvent choisir le cluster dans lequel ils sont

- Typiquement: cabinets médicaux

la localisation, l'âge et le sexe du médecin, les tarifs... peuvent influencer le choix des patients

- Autre exemple: commune d'habitation

Donner A & Klar N.
Design and analysis of cluster randomization trials in health research.
Arnold, London, 2000



Exemple (2): les individus dans un cluster sont affecté de manière similaire par un traitement

- Typiquement: patients opérés par un même chirurgien, en psychothérapie chez le même psychiatre...

dans ces cas l'issue peut être fortement liée aux caractéristiques des soignants et pas seulement à celles du patient

Donner A & Klar N, 2000



Exemple (3): liens privilégiés entre les individus du cluster

- Typiquement: familles

conséquences par exemples sur les possibilités de transmission de maladies infectieuses

Donner A & Klar N, 2000



Exemple (4): les individus exposés à une même condition interagissent entre eux

- Typiquement: thérapies de groupe

la dynamique au sein d'un groupe peut différer par rapport à celle dans un autre groupe

Donner A & Klar N, 2000



Conséquences

- Les réponses (ou mesures) des individus au sein d'un cluster seront plus similaires entre elles que si ces individus étaient dans des clusters différents
- Quand le recrutement se fait au niveau de clusters mais l'analyse au niveau d'individus, il faut en tenir compte dans le calcul de la taille de l'échantillon (qui devra être plus grand)



Coefficient de corrélation intraclasse (intracluster)

- Mesure le degré de similarité au sein de chaque cluster (corrélation de Pearson entre n'importe quelle paire de réponses au sein du cluster)
- Notation: ρ
- Utilisé pour calculer un indice de correction de la taille de l'échantillon



Coefficient de corrélation intraclasse (intracluster)

$$\rho = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}$$

σ_b^2 = variance inter-cluster

σ_w^2 = variance intra-cluster

Si $\rho > 0$: la variation entre les observations dans les différents clusters est plus grande que la variation au sein des clusters



Coefficient de corrélation intraclasse (intracluster)

- Le plus souvent <0.05 , mais peut être plus élevé
- A déterminer pour les issues principales d'une étude sur la base de données pilote ou dans la littérature:

Ukoumunne OC et al. Health Technology Assessment
1999; Vol 3: No 5

Adams G et al. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004;**57**(8):785-794



Facteur de correction pour le calcul de la taille d'un échantillon

- Notation: *Deff* (« *design effect* »)*
- Peut être calculé à partir de ρ selon la formule suivante:

$$Deff = 1 + \rho (m-1)$$

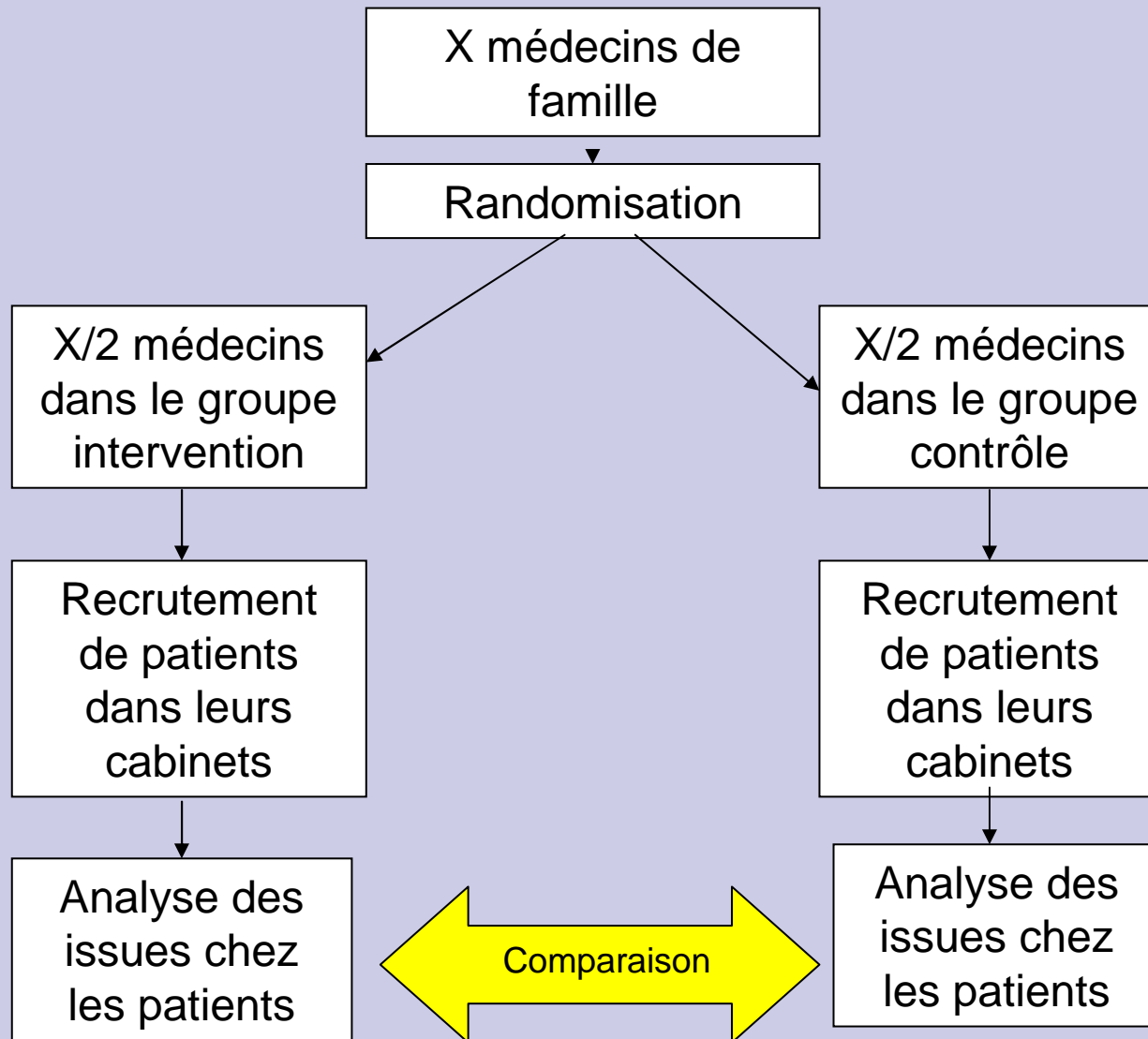
m = taille des clusters

*aussi appelé IF (« variance inflation factor »)

Exemple

- Calcul d'échantillon pour une étude randomisée en grappe dans des cabinets de médecine de premier recours







1. Calcul de la taille d'échantillon

« classique »

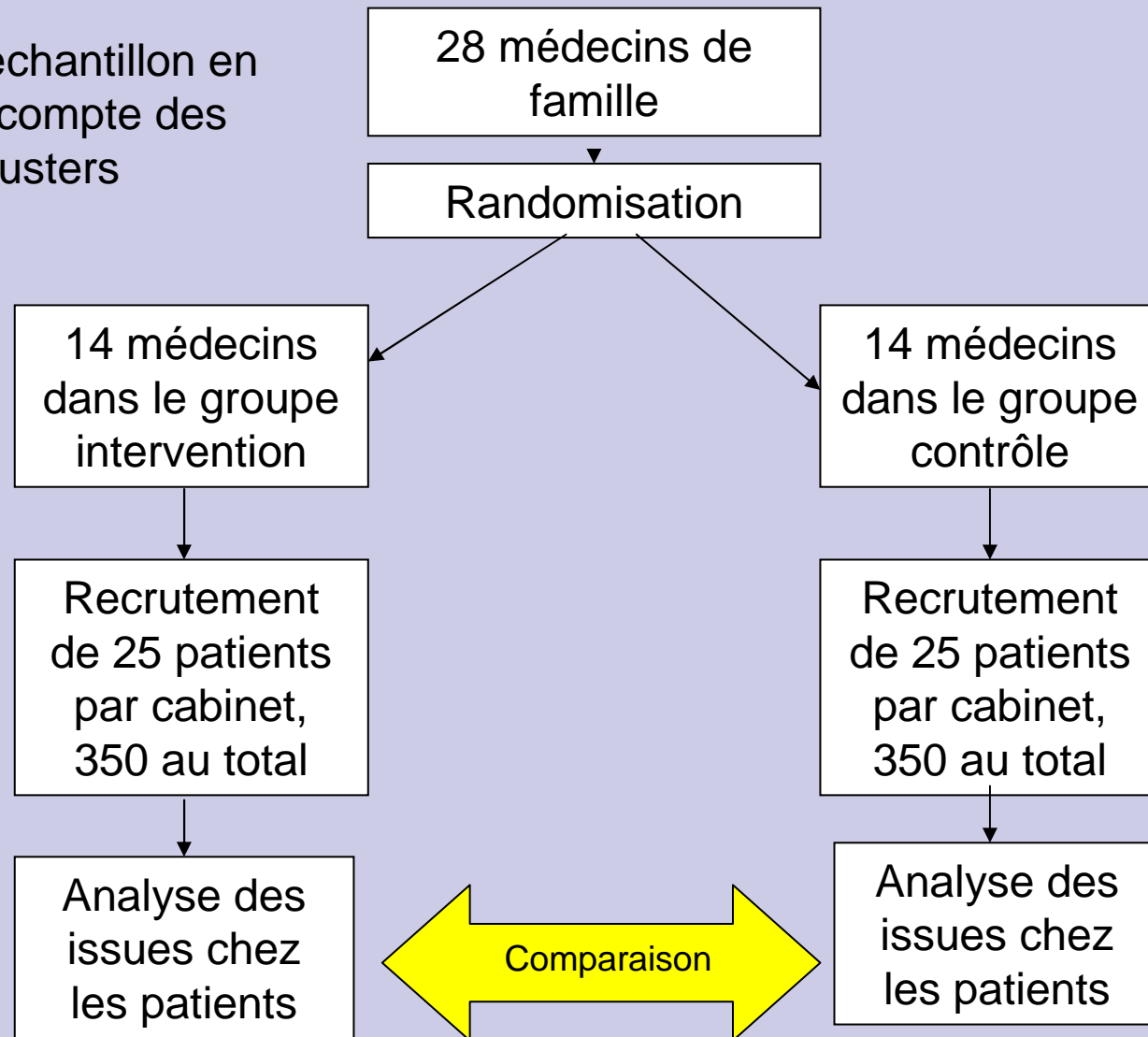
- Dans le cas présent nous avons déterminé qu'il nous faudrait 170 patients dans chaque groupe



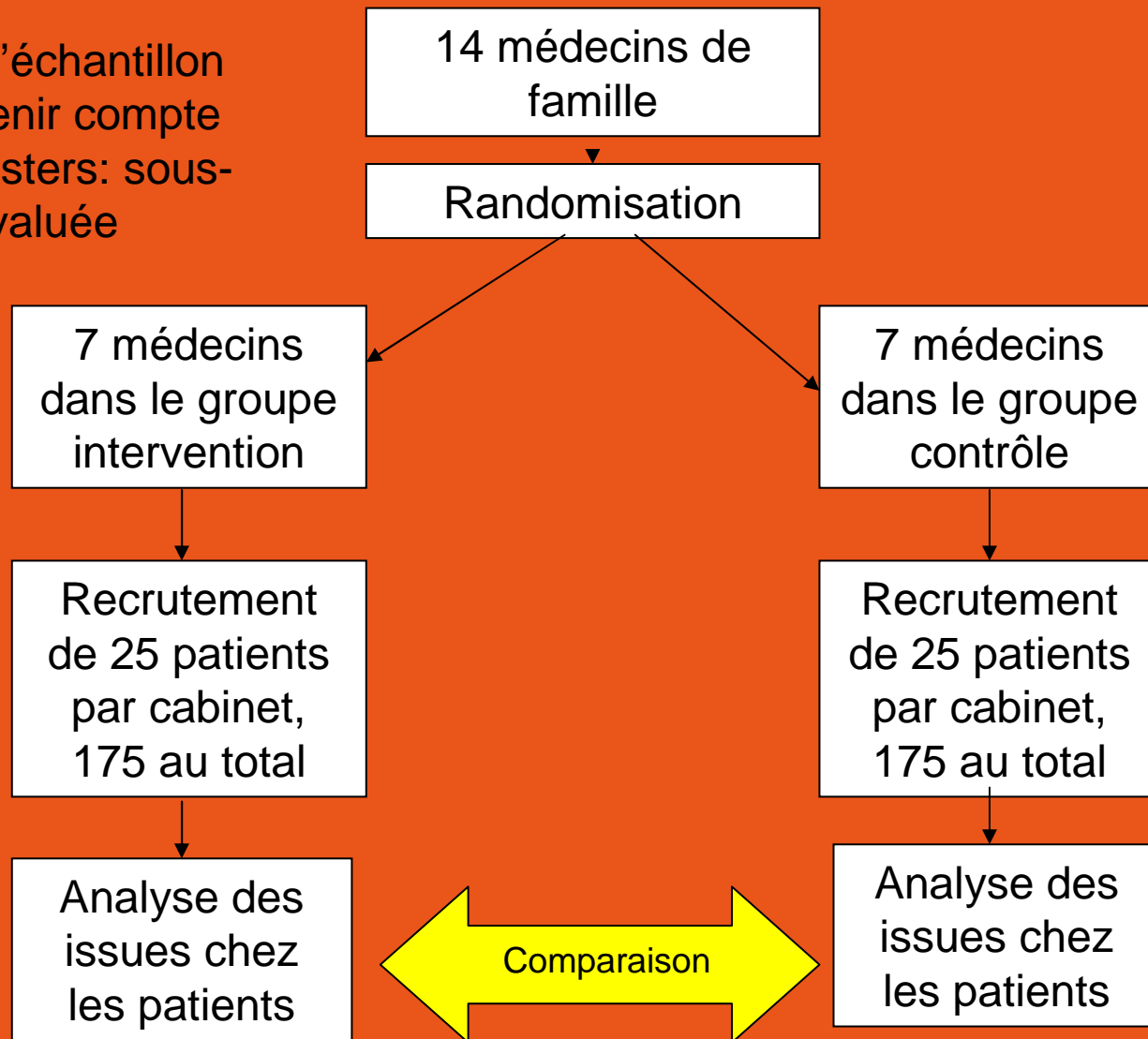
2. Calcul de la taille d'échantillon ajusté pour l'effet cluster

- Sur la base de l'étude pilote et de la littérature nous estimons que ρ pour les variables que nous voulons mesurer est ≤ 0.03
- $Deff = 1 + 0.03 (m-1)$
- Si $m = 25$, $Deff = 1.72$ et la taille de l'échantillon = **$1.72 * 340 = 585$**
- + 20% d'attrition = **700 patients**
- Avec 25 patients par médecins il faut **28 médecins** pour cette étude

Taille d'échantillon en
tenant compte des
clusters



Taille d'échantillon
sans tenir compte
des clusters: sous-
évaluée





Merci pour votre attention!

- Questions?

dagmar.haller-hester@hcuge.ch