

Faculté de Médecine d'Oran  
service de Biostatistique

*Statistique et  
Recherche clinique*

ANNEE UNIVERSITAIRE  
2015/2016

**Définition de la recherche dans le  
domaine de la santé**

- la recherche et les méthodes scientifiques peuvent être considérées comme une suite d'interrogations décisives conduisant à la découverte de faits ou d'informations qui améliore notre compréhension de la santé et des maladies de l'être humain.
- La recherche consiste en une quête des connaissances au moyen d'une enquête, d'une étude ou d'une expérimentation conduite avec application dans le but de découvrir et d'interpréter des connaissances nouvelles.

## Définition de la recherche dans le domaine de la santé

- La méthode scientifique est un ensemble systématique de procédures et de techniques relatives à l'exécution d'une étude ou d'une expérience destinée à acquérir de nouvelles connaissances.

---

## Catégories de recherches

## ***1. Recherche empirique, recherche théorique***

- Les recherches dans le domaine de la santé sont principalement de nature empirique, c'est-à-dire qu'elles sont fondées plus sur l'observation et l'expérience que sur la théorie et l'abstraction.
- La recherche épidémiologique, par exemple, dépend de la collecte systématique d'observations sur des phénomènes de santé spécifiques au sein d'une population définie.

## ***1. Recherche empirique, recherche théorique***

- De plus, même dans l'abstraction de modèles mathématiques, aucun progrès dans la compréhension de l'apparition et de la cause des maladies n'est possible sans comparer les constructions théoriques avec la situation réelle observée dans la population.
- La recherche empirique et la recherche théorique se complètent pour développer la compréhension des phénomènes, pour prévoir des événements futurs, ainsi que pour prévenir des événements nuisibles au bien-être général de la population concernée.

## *1. Recherche empirique*

- La recherche empirique dans le domaine de la santé peut être de nature qualitative ou quantitative

Pour la plus grande part, il s'agit de :

- l'identification de la population concernée, les caractéristiques (variables) des individus (unités) composant cette population,
- et enfin l'étude de la variabilité de ces caractéristiques parmi les individus de la population.

## *1. Recherche empirique*

- Ainsi, dans cette recherche empirique, la quantification est réalisée par trois procédures numériques associées :

- (a) mesure des variables ;
- (b) estimation des paramètres de la population (paramètres de la distribution de probabilité qui reflète la variabilité des observations dans la population) ;
- (c) test statistique des hypothèses, c'est-à-dire, estimation de la part contribué par le "hasard" dans les écarts observés parmi les individus ou les groupes étudiés.

## ***2. Recherche fondamentale, recherche appliquée.***

- Du point de vue fonctionnel, la recherche peut être fondamentale(ou pure) ou appliquée.

On considère d'habitude que la recherche fondamentale se consacre à l'acquisition des connaissances sans but défini d'utilité ou d'objet spécifique.

- La recherche appliquée découle d'un problème et elle est dirigée pour résoudre un problème existant.

Le bien-fondé et les bénéfices relatifs pour la société de la recherche fondamentale ou appliquée sont continuellement l'objet de controverses.

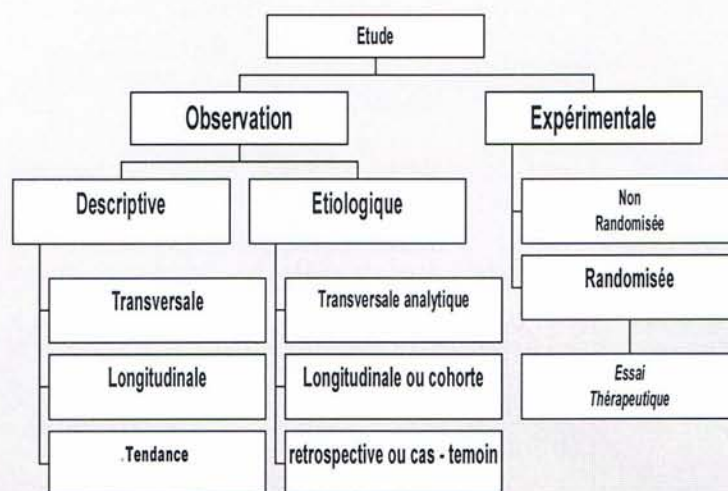
## ***2. Recherche fondamentale, recherche appliquée.***

- La recherche scientifique se consacrent directement à résoudre les problèmes relatifs à l'homme,
- Elle est la plus productive lorsqu'elle est entreprise librement ; en fait les plus grands progrès dans les sciences résultent de la recherche pure ou fondamentale.
- Il existe un équilibre entre les deux types de recherche.

## Les grands types d'études cliniques

- L'approche épidémiologique est fondée sur des principes statistiques dans l'élaboration du plan de recherche. Dans cette démarche, la recherche peut être subdivisée en recherche de type
  - basé sur l'observation,
  - et recherche de type expérimental.

### Types d'études:



## Types d'études:

### Les grands types de recherche

La plupart des études s'inscrivent dans l'une ou l'autre des catégories suivantes :

- **Thérapeutique** : mesurer l'efficacité d'un médicament, d'une intervention chirurgicale, d'une nouvelle méthode d'éducation des patients ou de toute autre procédure. Le protocole de choix est l'essai comparatif randomisé.
- **Diagnostic** : montrer si un nouveau test diagnostique est valide (peut-on lui faire confiance ?) et fiable (fournit-il les mêmes résultats lorsqu'on le répète ?). Le protocole de choix est l'enquête transversale (*cross sectional survey*) où sont mis en œuvre simultanément le nouveau test et l'examen de référence (*gold standard*).
- **Dépistage** : estimer la valeur d'un test qui peut être appliqué à une vaste population et qui détecte une maladie à un stade présymptomatique. La méthode de choix est l'enquête transversale (*cross sectional survey*).
- **Pronostic** : prévoir comment évoluera une maladie diagnostiquée chez un individu à un stade précoce. La méthodologie de choix est l'étude de cohortes longitudinale.
- **Etiologie** : démontrer si un agent présumé pathogène, une pollution environnemental par exemple, est effectivement impliqué dans la survenue d'une maladie. Le protocole de choix est soit l'étude de cohortes, soit l'étude cas-témoins, selon la fréquence de la maladie mais les séries de cas peuvent également apporter des informations cruciales.

## Deux grands types d'études

ETUDES NON  
EXPERIMENTALES

(études  
d'observation)

réalité observée telle qu'elle se  
présente

ETUDES  
EXPERIMENTALES

manipulation de  
l'exposition au facteur  
étudié pour ensuite  
observer l'effet

## Études d'observation

**Etudes  
Descriptives**

**Etudes  
Étiologiques**

## Etudes non expérimentales

---

- ETUDES DESCRIPTIVES

Décrivent un problème de santé dans une population ou un groupe d'individus et en établissent la fréquence selon certaines variables de personnes, de lieux et de temps.

- ETUDES A VISEE ETIOLOGIQUE

Cherchent à déterminer le rôle que peuvent jouer un ou des facteurs dans l'étiologie d'une maladie.



## Etudes descriptives

**TRANSVERSALES** : étude d'un problème de santé à un moment donné (génèrent des données de prévalence)

**LONGITUDINALES** : observation sur une période de temps (génèrent des données d'incidence)

## Etudes descriptives

**DE TENDANCE** : étude de la fréquence du phénomène à des périodes ou des dates données

(au moins trois moments différents)

BUT

Permet de dissocier  
trois caractéristiques liées au temps

Age, Période, Génération

## Etudes étiologiques.

- Cherchent à déterminer le rôle que peuvent jouer un ou des facteurs dans l'étiologie d'une maladie.

Existe -t-il une relation entre le facteur et la maladie ?

Quelle est l'intensité de cette relation ?

Elles peuvent être :

- **TRANSVERSALES**
- **LONGITUDINALES** : cohorte prospective
- **RÉTROSPECTIVE** : cas – témoin

## Deux grands types d'études

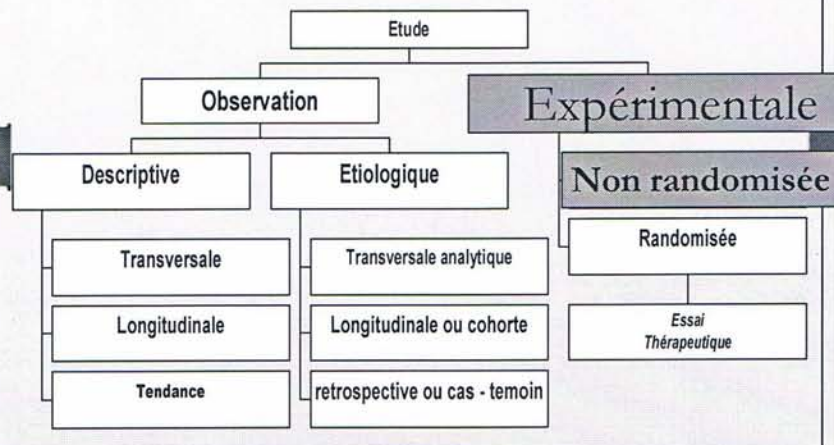
**ETUDES NON  
EXPERIMENTALES**  
(études d'observation)

réalité observée telle qu'elle se  
présente

**ETUDES  
EXPERIMENTALES**

manipulation de  
l'exposition au facteur  
étudié pour ensuite  
observer l'effet

## Types d'études: Récapitulatif



## Études non randomisées

La Constitution de deux groupes de population

Témoin et Étude est décidée par le chercheur

(pas de tirage au sort)

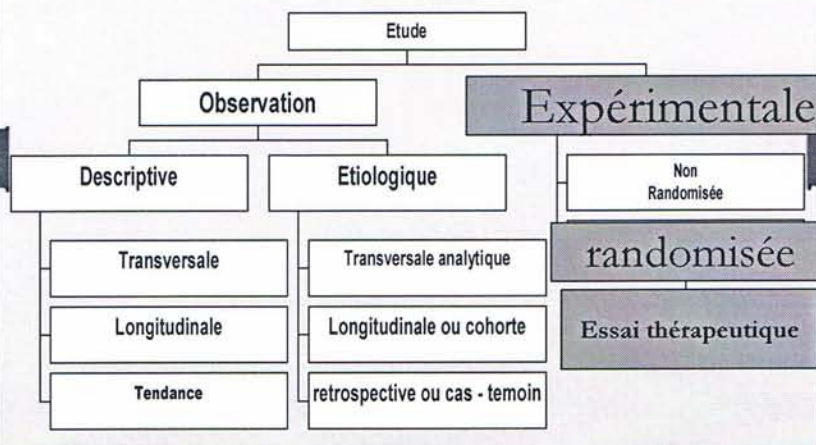
**Études d'évaluation  
de programmes de prévention**

## Etude non randomisée

### Etude quasi expérimentale

- Application d'une thérapeutique, un vaccin ou une intervention à un groupe d'individus choisis.
- Comparaison avec un groupe témoin.
- Jugement sur l'efficacité.

## Types d'études: Récapitulatif



## Études randomisées

---

Exemple type

**ESSAI THERAPEUTIQUE**

## Études expérimentales

- Evaluation en Santé

---

On distingue :

Les enquêtes expérimentales = étalon-or

→ IMPUTATION CAUSALE

Les enquêtes d'observation → PAS D'IMPUTATION  
CAUSALE au sens strict → Nécessité d'un Faisceau  
d'arguments

## Études expérimentales

Procédure expérimentale

- Intervention délibérément choisie
  - Appliquée à des sujets choisis
  - Selon une procédure déterminée par l'expérimentateur
  - Donc toujours prospective
  - Randomisée ou non
  - Procédure de choix en Recherche Clinique (recherche chez le sujet malade)

EX: évaluation de l'effet d'un nouveau traitement

- constitution de 2 groupes identiques en tout sauf en ce qui concerne le traitement à évaluer
- Par tirage au sort ou randomisation
- Ainsi une différence entre les deux groupes à l'issue de l'expérience pourra être imputée à la différence de traitement

## Études expérimentales

Cadre d'une démarche à visée évaluative

Manipulation  
du facteur  
à tester (évaluer)

*Contrôle total par observateur*

- Essais laboratoire
- Essais terrain

## Études expérimentales

Constitution de deux groupes d'individus

(intérêt de la comparaison)

- L'un concerné par le problème  
« Groupe d'Étude »
- L'autre non concerné par le problème  
« Groupe Témoin »

## Etudes expérimentales

Domaines d'utilisation :

- essais cliniques , essais thérapeutiques
- essais préventifs

Types :

- NON RANDOMISEES
- RANDOMISEES

## Études randomisées

Constitution de deux groupes d'individus

---

Témoin et Étude

se fera aléatoirement

(randomisation)

selon les principes du sondage

## Études randomisées

En simple aveugle

---

(simple insu)

Affectation des individus  
par tirage au sort,  
sans que l'individu (le sache).



## Études randomisées En double aveugle

---

(double insu)

Affectation des individus  
par tirage au sort,  
chercheur et individus eux mêmes  
ne savent pas  
Qui est «Témoin» et qui est «Étude»?

## Études randomisées et contrôlées Ouvert

---

Insu n'est pas maintenu  
Comparaison d'une méthode chirurgicale  
vs une méthode médicale  
ou  
Comparaison d'une méthode de chimiothérapie  
vs une méthode de radiothérapie

## Types d'essais cliniques

---

Selon la chronologie du développement

## Essais thérapeutiques : Phases

### Étape pré clinique

- Expérimentation en laboratoire
- In vitro
- In vivo chez l'animal

### Étape clinique

- Phase I
- Phase II
- Phase III
- Phase IV

## Étape clinique

**Phase I:**

- 1ère administration
- petit groupe de volontaires sains
  - ↳ Tolérance.

**Phase II:**

- petit groupe de volontaires malades
  - ↳ Modalités thérapeutiques efficaces (dose effet).

**Phase III:**

- grand groupe de volontaires malades
  - ↳ Efficacité thérapeutique.

**Phase IV:**

- En population, utilisation courante (détection des effets indésirables)

## En pratique...

Dépend du plan d'expérience :

- Nombre de groupes
- Indépendant / Apparié (patient propre témoin)

Dépend du critère de jugement principal

- Numérique
- Binaire
- Survie

De 2 risques :

- $\alpha$  : risque de première espèce : généralement 5%, de première espèce  
seuil de signification du test ( $H_0$  est rejeté)
- $\beta$  : risque de seconde espèce : inférieur à 20%, de deuxième espèce  
( $H_0$  est retenu)

## Notion de puissance d'un test

En pratique :

- on fixe  $\alpha = 5\%$
- on se donne  $\Delta$  sur critères cliniques  $\Delta = k/2$ ;  $k$  étant l'amplitude
- on estime  $S$  (étude pilote / littérature);  $S$  étant l'écart type.
- on calcule le nombre de sujets nécessaires pour assurer  $\beta = 10\%$  ou  $20\%$  (puissance  $> 80\%$ )

⇒ Formules, tables

## Nombre de cas à inclure dans l'étude?

On a déterminé le problème clinique

On a déterminé le(s) critère(s) de jugement

On a défini le plan d'expérience adapté



**Pb : combien, doit-on inclure de patients pour répondre correctement à l'hypothèse posée ?**

On utilise un test statistique ⇒ Notion de puissance

## Nombre de cas à inclure dans l'étude?

- $N = \frac{1.96^2 S^2}{\Delta^2}$  (caractère quantitatif)

$$\frac{1.96^2 S^2}{\Delta^2}$$

- $N = \frac{1.96^2 PQ}{\Delta P^2}$  (caractère qualitatif)

$$\frac{1.96^2 PQ}{\Delta P^2}$$

## Comment déterminer la taille d'un échantillon ?

- Cas enquête descriptive : nécessite de
  - connaître une approximation du caractère étudié
  - définir un degré de précision

- Calcul du nombre de sujets nécessaires:

$$n = \frac{\varepsilon^2 p(1-p)}{i^2}$$

$\varepsilon$  : écart-réduit (1,96 pour un risque  $\alpha = 5\%$ )

$p$  : prévalence estimée

$i$  : précision désirée

### ▪Exemple:

On veut évaluer la fréquence d'une maladie dans une population . Cette fréquence est de l'ordre de 10%

La précision désirée est de 2% pour un risque consenti de 5%.

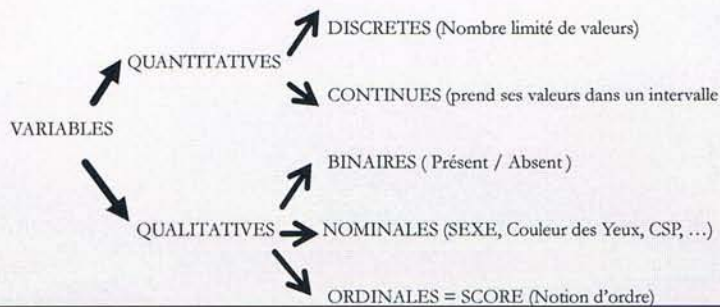
$$n = [(1,96)^2(0,1)(1 - 0,1)] / (0,02)^2 = 900$$

Le nombre de sujets nécessaires dans notre cas est de **900** individus.

## Méthodes statistiques : définitions générales

- **INDIVIDU** : Objet sur lequel un ou plusieurs caractères peuvent être observés.
- **POPULATION** : Ensemble des individus pris en considération.
- **VARIABLE** : Propriété servant à distinguer les individus d'une population.

Un caractère peut être **qualitatif (attribut)** ou **quantitatif**.



## La statistique descriptive

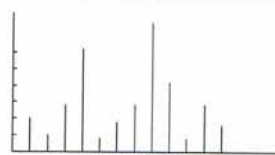
BUTS :

- Contrôle de qualité des données, descriptifs simples (moyennes, ...).
- Synthétiser, résumer, structurer l'information contenue dans les données.
- Mettre en évidence des propriétés de l'échantillon.
- Suggérer des hypothèses.

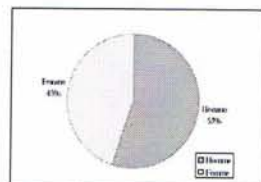
Analyses univariées : moyennes, histogramme, box-plot, fréquences, ...

Analyses multivariées = Analyse des Données. Permet de traiter des données multidimensionnelles.

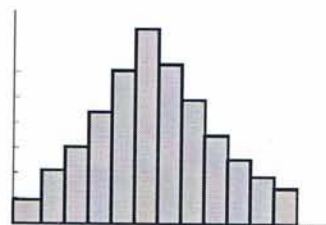
## Représentations graphiques



VARIABLES DISCRETES



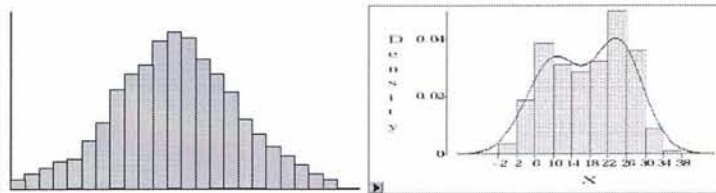
VARIABLES QUALITATIVES



VARIABLES CONTINUES

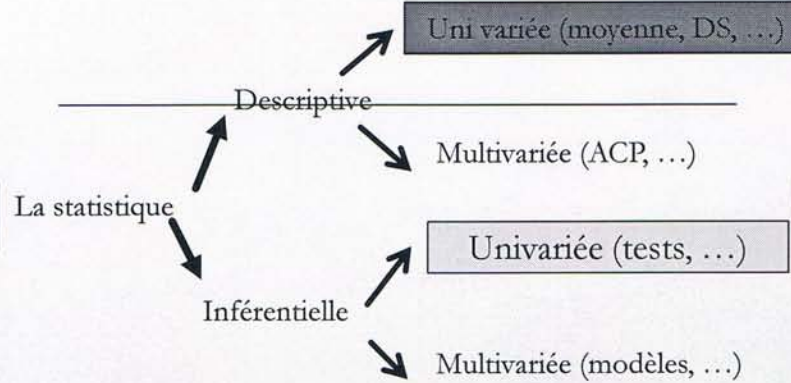
## Distribution d'un paramètre (loi)

- ◆ Différentes formes observables



- ◆ Modélisation de la distribution : Hypothèse de loi

## Les méthodes statistiques





## Analyse descriptive univariée

### 3 Objectifs :

Contrôle des données : Fréquences et Box-plots

Calcul des statistiques descriptives : moyenne, écart type

Présentation des résultats :

- **Moyenne  $\pm$  Déviation standard**
- **Fréquence avec Intervalle de confiance**

## Paramètres statistiques de base

◆ Moyenne :  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

◆ Variance estimée:  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$   
□

◆ Déviation standard : racine carrée de la variance

◆ Min, Max, Médiane, Quartiles, Centiles

## Intervalles de confiance à 95%

◆ d'une moyenne :  
quelque soit la loi de X, si  $n > 30$   $\Rightarrow \bar{x} \pm \frac{1.96 \sigma}{\sqrt{n}}$

◆ d'une fréquence  
si  $np, nq > 5$   $\Rightarrow p \pm 1.96 \sqrt{\frac{p(1-p)}{n}}$

## La statistique inférentielle univariée

### BUT :

Valider ou infirmer des hypothèses a priori ou formulées après une phase exploratoire.

Utilisation de tests statistiques se référant à des modèles probabilistes.

### EXEMPLES :

- Comparaison de moyennes
- ANOVA (+ + + !!!)
- Comparaison de fréquences
- Tests de lois

## Comparaison de groupes

- ◆ Choix du test statistique – Dépend de :
  - La nature de la variable
  - Du nombre de groupes
  - De la taille des groupes
- ◆ Cas des variables binaires ou qualitatives : Test du  $\chi^2$  ou Fisher exact
- Cas des variables numériques : Ecart-réduit, Student, ANOVA

## Mesures de signification statistique: valeur $p$

- La probabilité, dans une hypothèse nulle, de données aussi extrêmes ou plus extrêmes que celles observées dans une étude
- Fixée par convention à 0,05
- Equivalente à 5%
- La différence est significative si la valeur de  $p$  est inférieure à 0,05 ( $< 0,05$ )
- Signifie qu'il existe un risque de moins de 5% que le résultat obtenu est dû au hasard, ou
- Certain à 95% que le résultat obtenu par l'intervention est exact

## Interprétation d'un test.

Le résultat d'un test bilatéral montre que  $U_0 = 4,5$ .

La table suivante donne les probabilités  $\alpha$  pour que  $U$  soit supérieur à 2, 3, 4 et 5 :

U	2(1.96)	3	4	5
$\alpha$	0,05	0,02	0,01	0,001

La valeur  $U_0 = 4,5$  est supérieure à la valeur seuil  $U_{5\%} = 2$ . on rejette donc  $H_0$ .

La valeur immédiatement inférieure à  $U_0$  est 4 soit  $U_{1\%}$ . On conclut donc on acceptant  $H_1$  avec un degré de signification  $p < 0,01$

## CHOIX DES TESTS STATISTIQUES

OBJECTIF PRINCIPAL	OBJECTIF INTERMEDIAIRE	DEMARCHE
Comparer 2 variables qualitatives	Comparer 2 pourcentages -Test de conformité ( $P, P_0$ ) -Test d'homogénéité ( $P_1, P_2$ )	- Appliquer le test de l'écart réduit $\varepsilon$ des %
Comparer 2 ou plusieurs distri. quali.	Comparer 2 distrib. ou plus -Test de conformité ( $D_0, D_1$ ) -Test d'homogénéité ( $D_1, D_2$ )	- Appliquer le test du Khi-deux $\chi^2$
Comparer 2 variables quantitatives	Comparer 2 moyennes -Test de conformité ( $\bar{X}, \mu$ ) -Test d'homogénéité ( $\bar{X}_1, \bar{X}_2$ )	- Appliquer le test de l'écart réduit $\varepsilon$ des moyennes
	Comparer 2 variances -comparer la dispersion -vérifier l'égalité des variances.	- Appliquer le test de Fischer- F
Comparer + de 2 varia. quantitatives	Comparer plusieurs moyennes (ANOVA)	- Appliquer le test de F Analyse de la variance
Etudier la corrélat. entre 2 vari. quant.	Vérifier l'existence ou l'absence de la relation	- Appliquer le test du coeffici. de corr. R