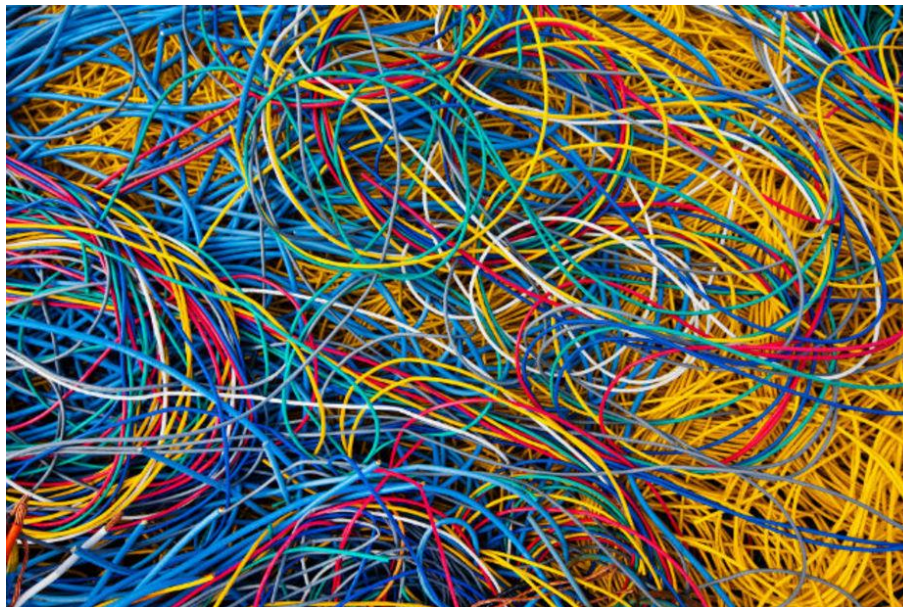


Recherche 3 en science infirmière

# Introduction aux tests statistiques et aux tailles d'effets

Pr Swann Pichon - Haute Ecole de Santé de Genève HES-SO  
Chercheur associé Université de Genève FPSE - Campus Biotech



h e d s

Haute école de santé  
Genève

HEdS  
Haute école de santé  
de Genève

47, av. de Champel  
1206 Genève  
+41 22 558 54 11

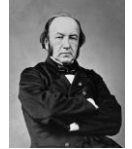
info.heds@hesge.ch  
www.hesge.ch/heds

Hes-so GENEVE  
Haute école spécialisée  
de Suisse occidentale

09 Janvier 2023

# Résumé des objectifs du dernier cours sur la démarche expérimentale

1. Pouvoir situer les origines historiques de la méthode expérimentale, et les origines de la méthode statistique qui y est associée
2. Pouvoir expliquer la différence **fondamentale** entre un lien de corrélation et un lien de causalité entre deux variables, et en quoi la démarche expérimentale permet de les différencier en mettant à l'épreuve l'hypothèse d'un lien de causalité entre les variables manipulées (VI) et la variable mesurée (VD)
3. Contraster les avantages et inconvénients de la démarche corrélationnelle et de la démarche expérimentale
4. Expliquer ce qui permet, dans la méthode expérimentale, de réduire les différences potentielles entre les participants et les groupes comparés
5. Expliquer la notion d'effet principal et d'interaction entre deux variables indépendantes que le devis expérimental fait varier
6. Identifier les différents types de devis expérimentaux et les deux grand types de mesures (inter versus intra-sujets)



# Plan et objectifs pédagogiques du cours



L'objectif pédagogique général du cours est de donner **une vue d'ensemble** de la notion de test statistique et des hypothèses qu'ils permettent de tester.

Nb: Notez que des notions comme les conditions d'applicabilité du test, le calcul et la distribution de la statistique, ou encore les notions de tests paramétriques et non-paramétriques sont des notions importantes non abordées dans ce cours

- Pouvoir expliquer la logique globale d'un test d'hypothèse, sans mathématiques
- Pouvoir expliquer le but d'un test d'association (encore appelé test de corrélation) et le but d'un test de différence de moyennes
  1. Association: Comprendre la logique du test du  $\chi^2$  (**Chi2**)
  2. Association: Comprendre la logique du test de **corrélation linéaire**
  3. Association: Comprendre la logique du test de **régression linéaire**
  4. Différence de moyennes : Comprendre la logique du **test de Student** et de **l'ANOVA**
- Pouvoir expliquer la notion de taille d'effet et son lien par rapport à la pertinence clinique d'une intervention

# Un premier mot avant de commencer



- L'objectif de ce cours était de vous faire comprendre la logique des principaux test statistiques et les inférences qu'elles permettent de tester
- Nous ne nous attendons pas à ce que vous reteniez tous les détails présentés aujourd'hui, ni à savoir appliquer ces tests concrètement, mais au moins à les lire et les comprendre
- **Nous vous encourageons à poser toutes les questions qui pourront vous aider et aider vos camarades à mieux comprendre les objectifs abordés dans ce cours**

# Les statistiques descriptives et différentielles - Définitions

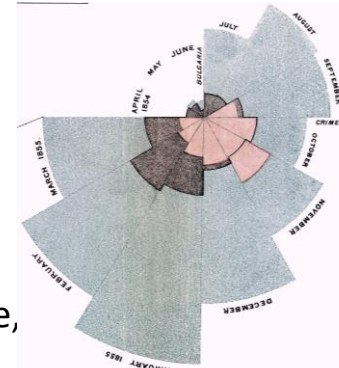
- La statistique peut se définir comme l'art de collecter, d'analyser, de visualiser et d'interpréter des « données » pour évaluer la « fiabilité » des décisions fondées sur ces données (i.e. existe-t-il un lien entre deux variables observées? ou une différence de moyenne entre la valeur d'une variable mesurée dans le cadre de 2 conditions d'une étude expérimentale, ou une différence entre deux échantillons observés?). On distingue deux types de statistiques complémentaires



F. Nightingale

- Les statistiques **descriptives**

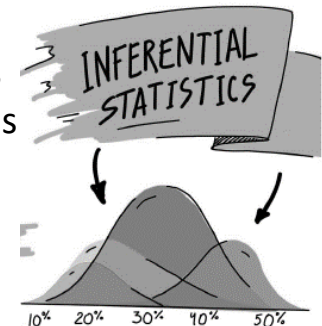
- Sont des méthodes pour **décrire** la tendance centrale (moyenne, médiane) d'une variable, sa fréquence, sa dispersion (déviatoin standard, IC95% ...)
- Représenter** ces tendances et ces distributions via des graphiques (boites à moustache, histogrammes, graphiques à barres, nuages de points ...)



Mortalité Hôpital Scutari

- Les statistiques **inférentielles** (aka les tests)

- Sont des méthodes probabilistes utilisées (par ex.) pour estimer l'existence de corrélations ou de différences de moyennes entre deux variables
- Ces tests sont utilisés car on a jamais accès à une population dans son ensemble, mais toujours à un sous-échantillon. Le but des tests statistiques est d'étendre les propriétés de l'échantillon à celui de la population duquel il est issu.
- Ces tests sont faits en acceptant une certaine probabilité d'erreur => la valeur de p (probabilité de réaliser un faux positif)

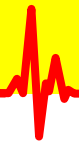




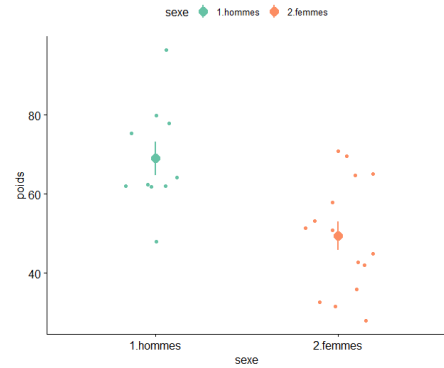
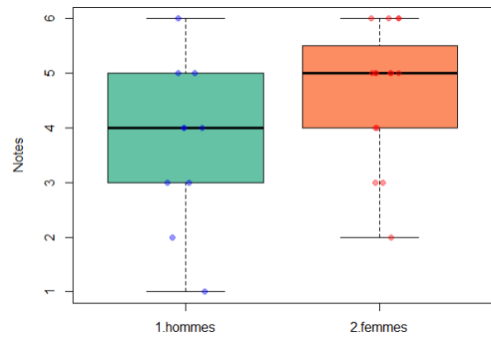
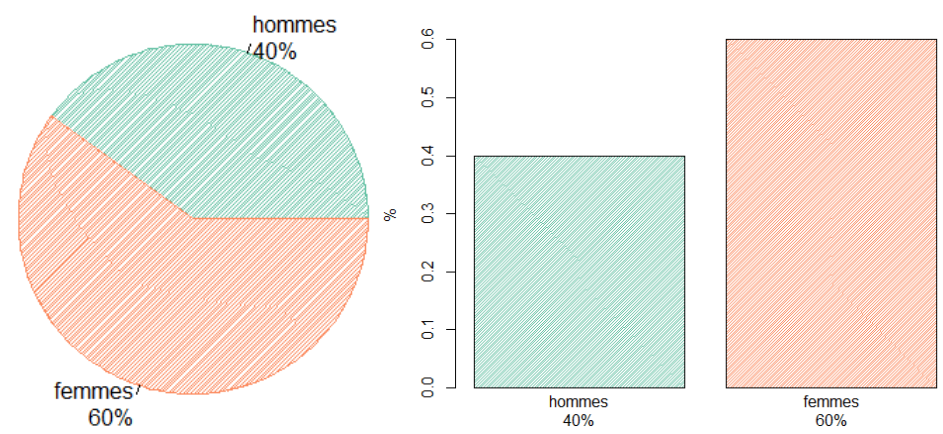
# Statistiques **descriptives** et statistiques **inférentielles**

<p>Echelle de mesure</p>	<p>Décrire : tendance centrale</p>	<p>Il p gra</p>		<p>s de es entre ennes</p>	<p>Tests de différences entre +2 moyennes</p>
<p>Catégorielle</p>	<p>Mode %, effectif</p>	<p>Car ou b fré</p>		<p>appariés opariés</p>	<p>NA = non-appariés A = appariés</p>
<p>Ordinale</p>	<p>Médiane %, effectif</p>	<p>B mc</p>		<p>arré/Chi- e (<math>\chi^2</math>) arré de emar Chi-square</p>	<p>NA: Chi-carré/Chi- square (<math>\chi^2</math>) A: Cochran (Q)</p>
<p>Continue</p>	<p>Moyenne</p>	<p>E</p>		<p>Test de Whitney- on (U)</p>	<p>NA: Kruskal-Wallis (H) A: Friedman (Q)</p>

# Statistiques descriptives



Type de Variable	Décrire la tendance centrale dans un texte (dispersion)	Illustrer la tendance par un graphique
Catégorielle (sexe: ♀:10, ♂:15)	% et effectif	i.e. Camembert ou barres de fréquences

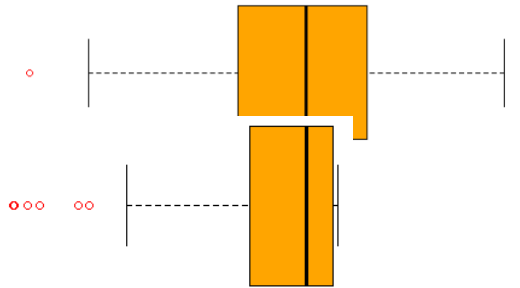
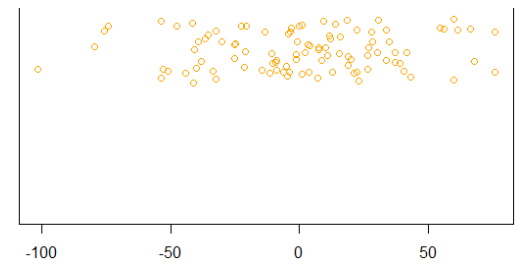
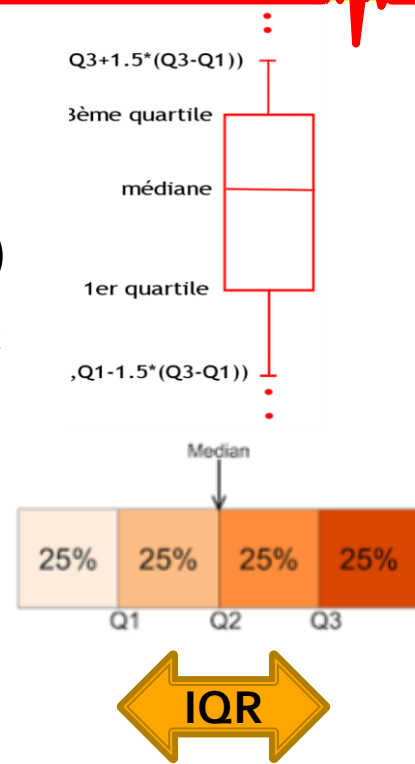


Boîtes à moustache

Moyennes

# Illustrer une médiane et les valeurs extrêmes d'une distribution à l'aide d'une boîte à moustache (BAM)

- Une BAM permet de visualiser en un seul graphique la médiane, les quartiles et les valeurs extrêmes de l'échantillon
- Souvent utilisée pour illustrer des données discrètes mais aussi continues car pratique pour visualiser les valeurs extrêmes (outliers)
  - Quartiles: Les 3 valeurs (Q1, Q2, Q3) qui divisent un jeu de données en 4 jeux de même taille
  - Médiane: la valeur Q2 qui divise un jeu de données en 2 jeux de même taille
  - Distance interquartile (**IQR**) =  $Q3 - Q1$
  - Les valeurs extrêmes d'une BAM sont  $< (Q1 - 1.5 * IQR)$  ou  $> (Q3 + 1.5 * IQR)$



Distribution de l'échantillon (n=100)

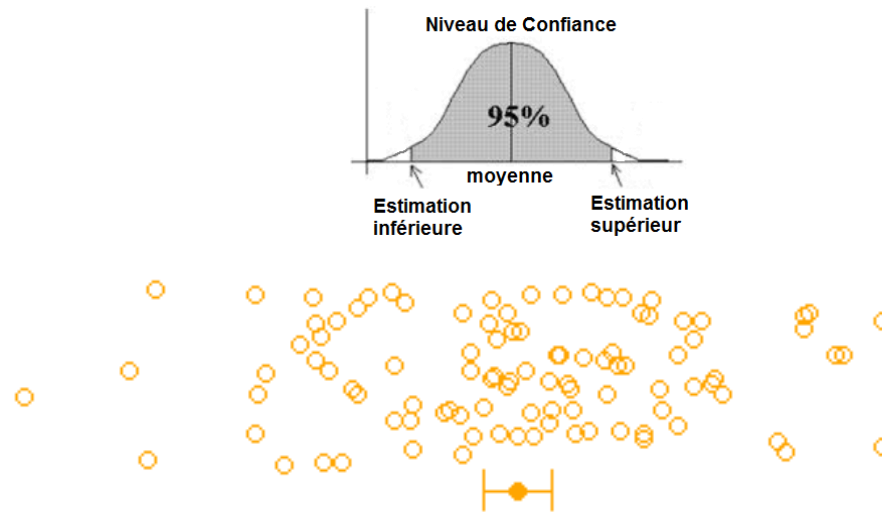
Distribution non problématique (n=100)

Autre distribution possiblement problématique



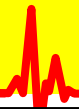
# Illustrer une moyenne et l'intervalle de confiance liée à son estimation

- Lorsque l'on cherche à estimer une moyenne au sein d'un *échantillon*, il est important de l'accompagner d'un indicateur de dispersion comme l'intervalle de confiance à 95% (IC95) qui renseigne sur la précision de son estimation
- L'intervalle de confiance à 95% est l'intervalle de valeur qui a 95% de chance de contenir la vraie valeur de la moyenne dans la *population* cible



- Exemple1: La réduction du poids suite au régime A est de -0.5 Kg [-1.2 +1.1]
- Exemple2: La réduction du poids suite au régime B est de -1.6 Kg [-2.3 -1.3]

# Statistiques inférentielles – Tests d'association et Tests de différences de moyennes



Type de Variable	Tests d'association (paramètre du test)		Tests de différence de moyennes (paramètre du test) NA = non-appariés; A = appariés	
	Lien ou <b>corrélation</b> entre 2 variables	Régression entre 2 variables	Entre 2 moyennes	Entre +2 moyennes
Catégorielle (sexe: ♀:10, ♂:15)	<b>Test Chi-carré (<math>\chi^2</math>)</b> <b>Test Phi (<math>\phi</math>)</b> <b>Test exact de Fisher (F)</b>	Régression logistique  (OR =>odds ratio; RR=>risque relatif)	NA: Chi-carré ( $\chi^2$ )  A: Chi-carré de McNemar ( $\chi^2$ )	NA: Chi-carré ( $\chi^2$ )  A: Cochran (Q)
Discrète (note: [1,2,3,4,5,6])	<b>Rho Spearman (<math>\rho</math>)</b> <b>Tau de Kendall (<math>\tau</math>)</b>	Régression linéaire (R ou R <sup>2</sup> et b ou $\beta$ )	NA et A: Test de Mann-Whitney-Wilcoxon (U)	NA: Kruskal-Wallis (H)  A: Friedman (Q)
Continue (poids: 70,3 kg)	<b>Corrélation de Pearson (r)</b>		NA et A: Test T de Student (t)	NA: ANOVA (F)  A: ANOVA à mesures répétées (F)

# Grille des tests statistiques ordonnés par type de groupes comparés (dépendant versus indépendant)



Groupes indépendants = Groupes non-appariés (NA) = individus différents dans chaque groupe  
 Groupes dépendants = Groupes appariés (A) = mêmes individus mesurés dans chaque groupe

Level of measurement of dependent variable	Group Comparisons: Number of groups (the independent variable)				Correlational analyses (To examine relationship strength)
	2 Groups		3+ Groups		
	Independent Groups Tests	Dependent Groups Tests	Independent Groups Tests	Dependent Groups Tests	
Nominal (Categorical)	$\chi^2$ p. 401 (or Fisher's exact test) p. 402	McNemar's test  p. 402	$\chi^2$  p. 401	Cochran's Q	Phi coefficient (dichotomous) or Cramér's V (not restricted to dichotomous) p. 403
Ordinal (Rank)	Mann-Whitney Test p. 396	Wilcoxon signed ranks test p. 396	Kruskal-Wallis H test p. 400	Friedman's test p. 400	Spearman's rho (or Kendall's tau) pp. 403
Interval or Ratio (Continuous)*	Independent group <i>t</i> test pp. 394-395	Paired <i>t</i> test p. 396	ANOVA pp. 396-399	RM-ANOVA pp. 400	Pearson's <i>r</i>  p. 402
	Multifactor ANOVA for 2+ independent variables			p. 398	
	RM-ANOVA for 2+ groups x 2+ measurements over time				

\*For distributions that are markedly nonnormal or samples that are small, the nonparametric tests in the row above (for ordinal measures) may be needed.

# Quatre grandes familles de tests (le cas des variables continues a été considéré pour l'exemple)

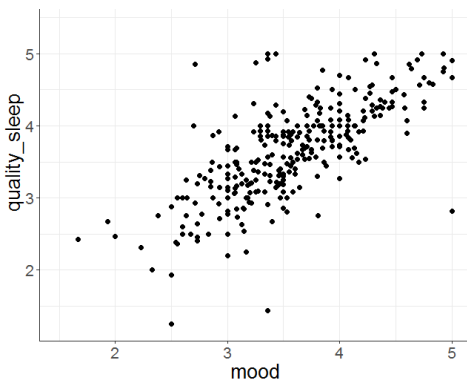
Tests de corrélation  
(lien)

Tests de régression  
(lien)

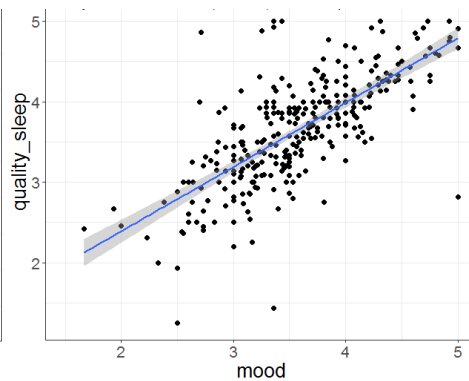
Tests de différences  
entre 2 moyennes

Tests de différences  
entre +2 moyennes

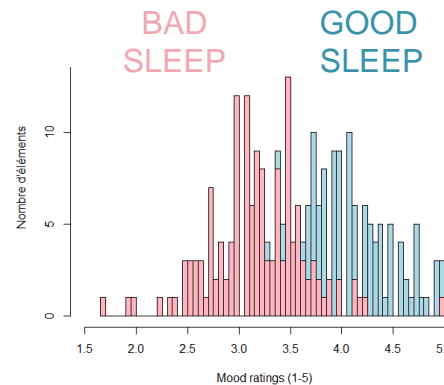
- Ce set de données représente les mesures d'humeur selon la qualité du sommeil chez ~200 hommes et femmes. Les 4 tests suivants et les graphiques associés ont été employé sur les mêmes données
  - Corrélation : la variable humeur et qualité du sommeil varient-elles ensemble? (oui, positivement)
  - Régression : existe-t-il une droite de pente non nulle qui approxime cette relation (oui, si la pente était plate, il n'y aurait pas de lien entre les variables).
  - T-test : existe-t-il une différence de moyenne entre l'humeur des bons et des mauvais dormeurs? (oui)
  - ANOVA : existe-t-il une différence de moyenne entre l'humeur des bons, moyens et mauvais dormeurs? (oui)



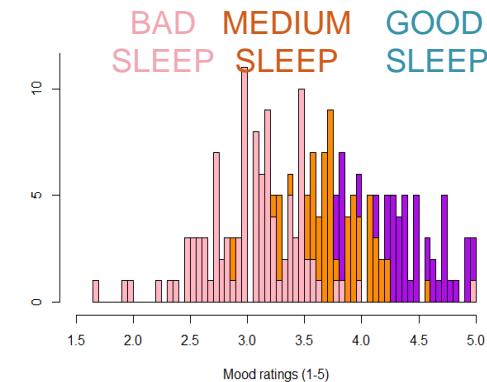
Corrélation de Pearson  
( $r$ )



Régression  
linéaire



Student t-test



ANOVA

# Principe général d'un test statistique et p-valeur



- Un test statistique est une procédure de décision entre deux hypothèses
  - $H_0$  : l'hypothèse nulle (pas d'effet)
  - $H_1$  : l'hypothèse alternative (un effet)
- Le principe de chaque test statistique consiste à
  1. Calculer un score statistique qui «résume les données testées»
  2. Puis décider si ce score penche plutôt en faveur de l'hypothèse nulle ou de l'hypothèse alternative, en acceptant un risque d'erreur de se tromper (la valeur  $p$ , qui exprime la probabilité de réaliser un faux positif)
- Chaque statistique estime un score particulier
  - Score **Chi2** lorsque l'on traite les fréquences d'un tableau de répartition entre deux variables
  - Score **r** de Pearson lorsque l'on teste une corrélation entre deux variables continues
  - Score **t** de Student si on teste une différence entre 2 moyennes
  - Score **F** de l'ANOVA si on teste une différence entre +2 moyennes



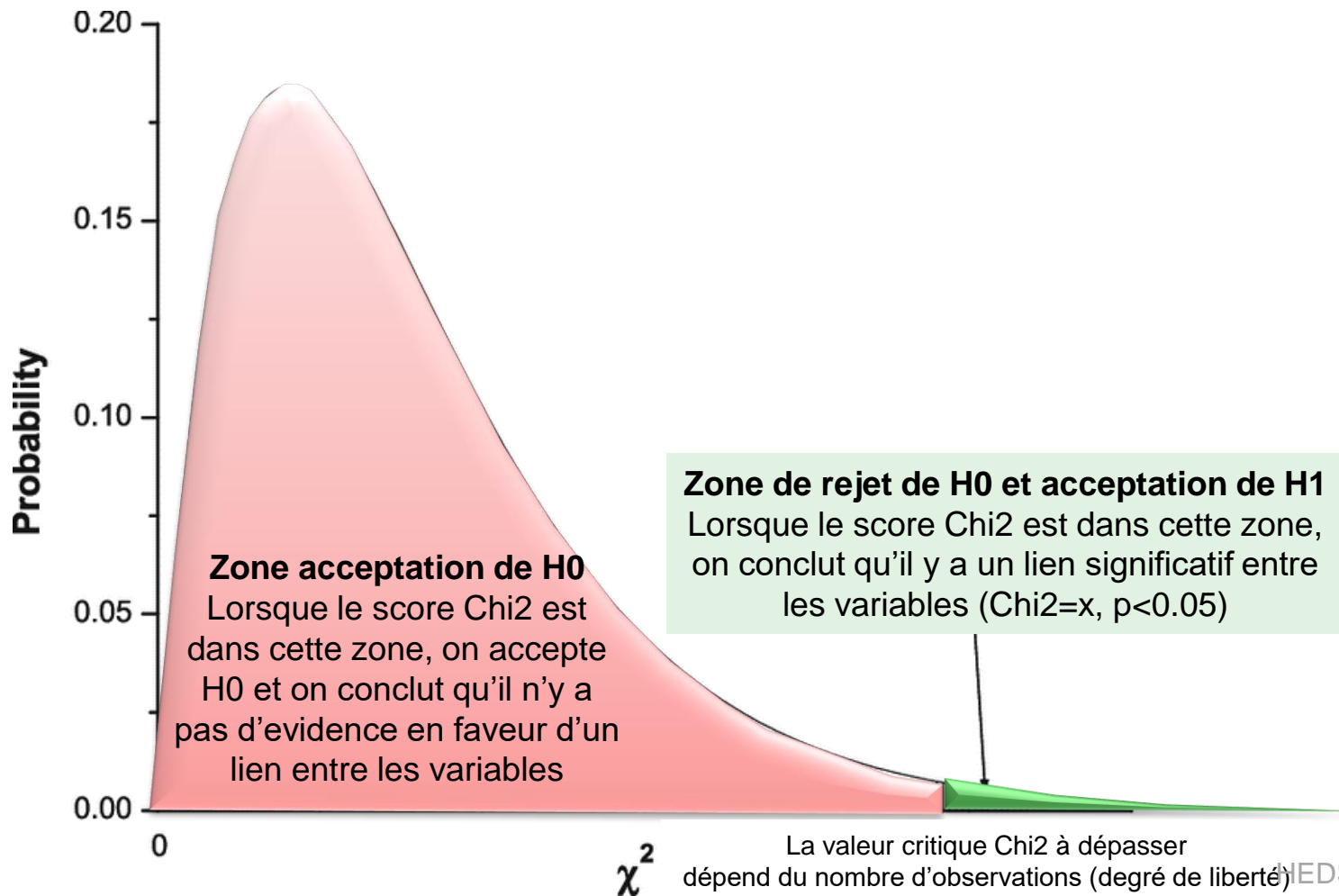
# Principe général d'un test statistique et p-valeur



- **La logique d'un test d'hypothèse** consiste toujours à faire un pari sur  $H_0$  pour conclure si l'hypothèse nulle est vraie ou improbable.
  - $H_0$ : l'hypothèse nulle est vraie => Il n'existe pas assez d'évidence dans les données pour rejeter  $H_0$
  - $H_1$ : l'hypothèse nulle est improbable => l'hypothèse alternative ( $H_1$ ) est vraie
- Cette décision de rejet de  $H_0$  est réalisée en acceptant un risque de se tromper en déclarant le test significatif (faire un faux positif)
  - => **c'est ce risque que quantifie la valeur p.**
  - Pour que le test soit déclaré significatif, le risque de se tromper doit être strictement inférieur au seuil alpha, qui par convention est 0.05 (5%)
- Par exemple, un test de corrélation entre la variable «qualité du sommeil» et le «niveau d'humeur» retourne une valeur  $r = 0.73$ , avec une valeur de  $p = 0.001$ .
  - Je peux donc déclarer ce test significatif car  $p < 0.05$
  - La valeur p m'informe que j'ai moins d'1 chance sur 1000 de me tromper en déclarant ce test significatif.
- **Chaque score statistique a une distribution particulière attendue sous  $H_0$** , qui est fonction du nombre d'observations dans l'échantillon (degré de liberté).

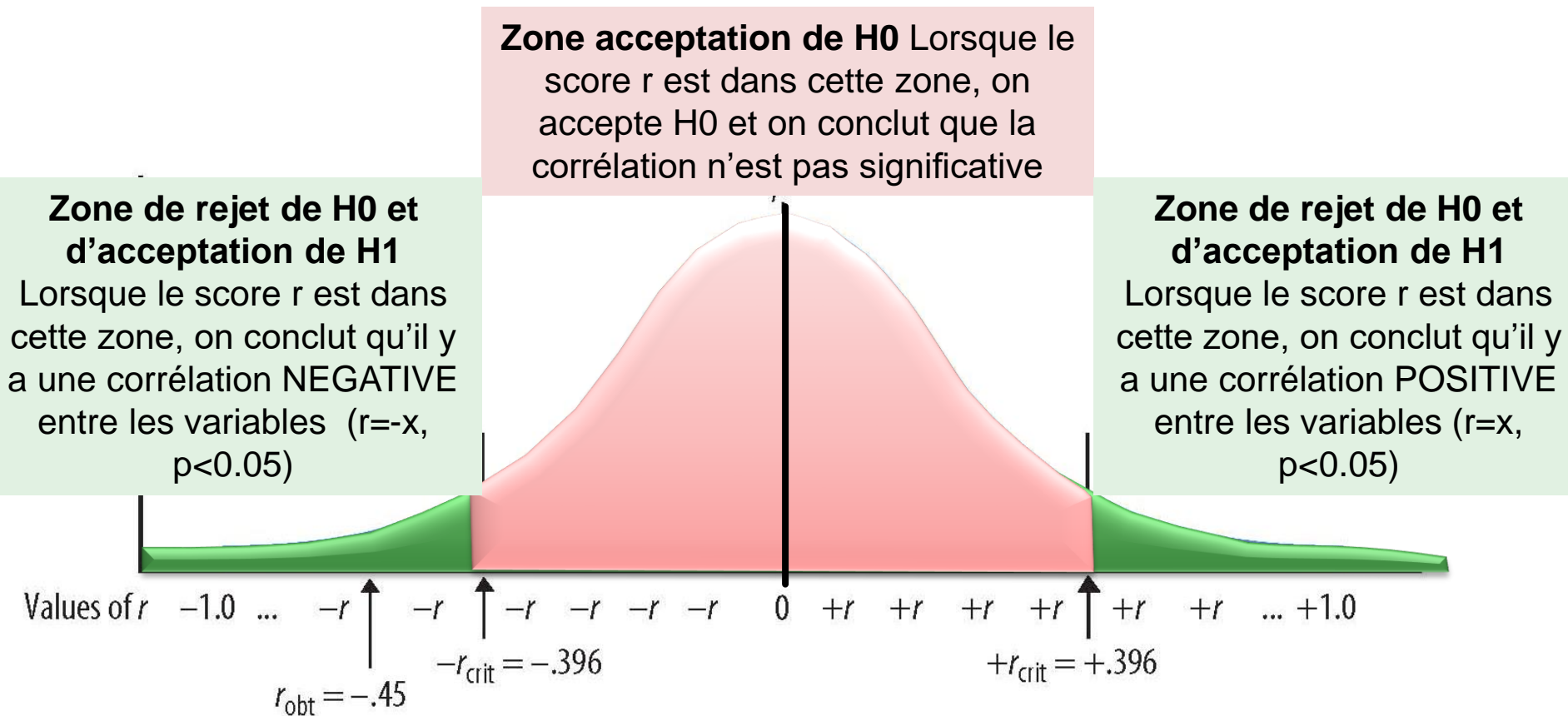
# Exemple de Distribution du Chi-carré sous H0 – test d'indépendance du Chi2

- Cette distribution, uniquement positive, permet de tester l'hypothèse d'un lien entre deux variables (ex: facteur de risque X patients)



# Exemple de Distribution du score r de Pearson sous H0 - Test bilatéral de corrélation de Pearson

- Cette distribution, centrée sur 0, permet de tester l'hypothèse qu'un score r positif ou négatif diffère ou non de 0



La valeur critique r à dépasser pour déclarer le test significatif dépend du nombre d'observations

# Tests d'hypothèses et inférences faites selon les tests (Chi2, corrélation, régression, T-test, ANOVA)

## ■ H0 = Hypothèse nulle ~ évidence insuffisante en faveur d'un effet ( $p \geq 0.05$ )

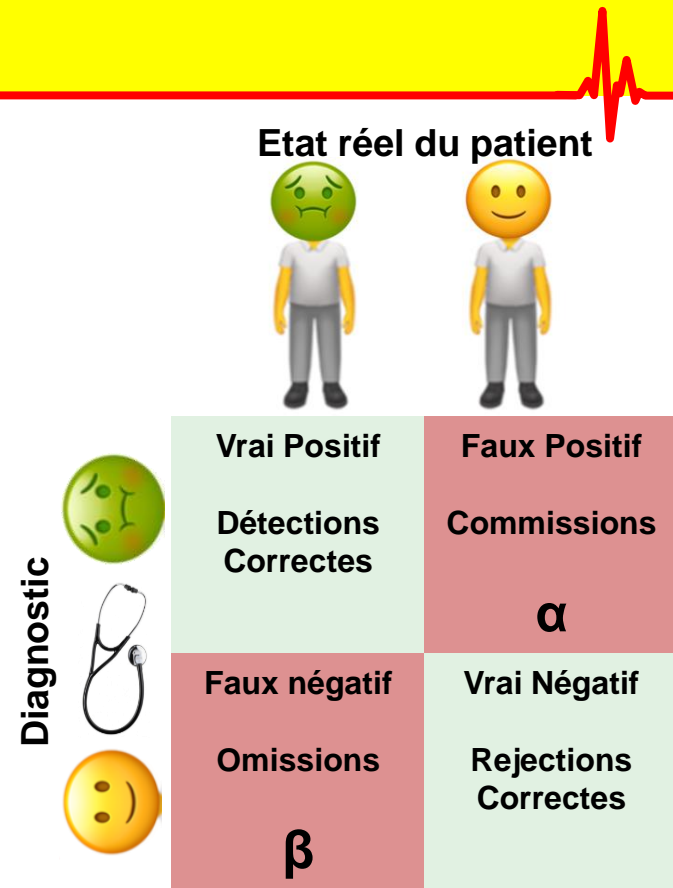
- **Chi2 ( $\chi^2$ ):** Le facteur d'exposition n'a pas d'effet significatif sur l'outcome clinique
- **Corrélation:** le coefficient de corrélation n'est pas significativement différent de 0
- **Régression:** la pente (ou l'intercept) de la droite n'est pas significativement différente de 0
- **T-Test de Student:** La différence entre deux moyennes n'est pas signif. différente de 0
- **ANOVA:** Les n moyennes comparées ne sont pas signif. différentes les unes des autres

## ■ H1 = Hypothèse alternative ~ évidence en faveur d'un effet ( $p < 0.05$ )

- **Chi2:** Le facteur d'exposition est lié significativement à l'outcome clinique
- **Corrélation:** le coefficient de corrélation est signif. différent de 0
- **Régression:** la pente (ou l'intercept) de la droite est signif. différente de 0
- **T-Test de Student :** Les deux moyennes comparées sont signif. différentes
- **ANOVA:** Au moins une des n moyennes comparées est signif. différente des autres

# On distingue deux types d'erreurs en statistique, ainsi que deux types de risques

- L'erreur de **faux positif** consiste à déclarer un test significatif (je déclare qu'un non-patient est malade) alors qu'en fait il ne l'est pas réellement (erreur de commission)
- **La risque alpha ( $\alpha$ )**, désigne le risque de déclarer un **faux positif** lorsque l'on rejette  $H_0$ . C'est ce seuil alpha que l'on fixe par convention à 5% et que l'on utilise pour décider si un test est significatif ou non via la **valeur p**.
- L'erreur de **faux négatif** consiste à ne pas déclarer un test significatif (je déclare qu'un vrai patient n'est pas malade) alors qu'en fait il l'est réellement (erreur d'omission)
- **Le risque beta**, généralement de 20%, désigne le risque de faire un faux négatif, c'est-à-dire de ne pas détecter un effet s'il existe vraiment. Le risque beta est Important pour réaliser un **calcul de puissance statistique** qui vise à déterminer la taille minimale de l'échantillon que l'on doit collecter pour se donner une chance de 80% de détecter un effet s'il existe.





# Plan et objectifs pédagogiques du cours



L'objectif pédagogique général du cours est de donner **une vue d'ensemble** de la notion de test statistique et des hypothèses qu'ils permettent de tester.

Nb: Notez que des notions comme les conditions d'applicabilité du test, le calcul et la distribution de la statistique, ou encore les notions de tests paramétriques et non-paramétriques sont des notions importantes non abordées dans ce cours

- Pouvoir expliquer la logique globale d'un test d'hypothèse, sans mathématiques
- Pouvoir expliquer le but d'un test d'association (encore appelé test de corrélation) et le but d'un test de différence de moyennes
  1. Association: Comprendre la logique du test du  $\chi^2$  (**Chi2**)
  2. Association: Comprendre la logique du test de **corrélation linéaire**
  3. Association: Comprendre la logique du test de **régression linéaire**
  4. Différence de moyennes : Comprendre la logique du **test de Student** et de **l'ANOVA**
- Pouvoir expliquer la notion de taille d'effet et son lien par rapport à la pertinence clinique d'une intervention

# Quatre grandes familles de tests statistiques – Apprenons à distinguer les 4 tests pour les variables continues

Type de Variable	Tests d'association (paramètre du test)		Tests de différence de moyennes (paramètre du test) NA = non-appariés; A = appariés	
	Lien ou <b>corrélation</b> entre 2 variables	Régression entre 2 variables	Entre 2 moyennes	Entre +2 moyennes
Catégorielle (sexe: ♀:10, ♂:15)	<b>Test Chi-carré (<math>\chi^2</math>)</b> <b>Test Phi (<math>\phi</math>)</b> <b>Test exact de Fisher (F)</b>	Régression logistique  (OR =>odds ratio; RR=>risque relatif)	NA: Chi-carré ( $\chi^2$ )  A: Chi-carré de McNemar ( $\chi^2$ )	NA: Chi-carré ( $\chi^2$ )  A: Cochran (Q)
Discrète (note: [1,2,3,4,5,6])	<b>Rho Spearman (<math>\rho</math>)</b> <b>Tau de Kendall (<math>\tau</math>)</b>	Régression linéaire (R ou R <sup>2</sup> et b ou $\beta$ )	NA et A: Test de Mann-Whitney-Wilcoxon (U)	NA: Kruskal-Wallis (H)  A: Friedman (Q)
Continue (poids: 70,3 kg)	<b>Corrélation de Pearson (r)</b>		NA et A: Test T de Student (t)	NA: ANOVA (F)  A: ANOVA à mesures répétées (F)

# Le principe du test Chi<sup>2</sup> (Chi-square, $\chi^2$ )

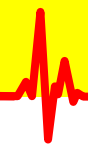


- De nombreuses études observationnelles cas-contrôle recueillent des données de variables catégorielles et résument les effectifs de ces catégories dans un **tableau de répartition des effectifs (ou de contingence)**.
- Par exemple, en épidémiologie, le test du Chi<sup>2</sup> est souvent utilisé dans les études de cas contrôle pour tester l'hypothèse d'un lien entre un facteur d'exposition (exposé/non-exposé) et un facteur de santé (ex: malade/non malade)
- La statistique du test du Chi<sup>2</sup> sert à tester l'hypothèse qu'il existe une association entre deux variables catégorielles.
  - H0: Les deux facteurs catégoriels sont indépendant (pas de relation entre les facteurs)
  - H1: Les deux facteurs sont dépendants (relation entre les facteurs)
- **La logique du test** vise à comparer les effectifs du tableau de répartition des cas observés, aux effectifs théoriques que l'on devrait observer s'il n'existait pas de relation entre les variables catégorielles.
- Enoncé autrement, ce test vise à mettre en évidence un pattern d'interaction entre deux variables catégorielles

Effectifs observés (nombre de cas)		
	Malades (Cases)	Non-malades (Controls)
Exposed	A	B
Non-exposed	C	D

Effectifs théoriques que l'on devrait observer s'il n'y avait pas de relation (Ho)		
	Malades (Cases)	Non-malades (Controls)
Exposed	A <sub>théo.</sub>	B <sub>théo.</sub>
Non-exposed	C <sub>théo.</sub>	D <sub>théo.</sub>

# Un exemple récent – le statut tabagique est-il lié à la sévérité de la COVID-19?



**TABLE 1. Reported outcomes among COVID-19 patients of all ages, by hospitalization status, underlying health condition, and risk factor for severe outcome from respiratory infection — United States, February 12–March 28, 2020**

Underlying health condition/Risk factor for severe outcomes from respiratory infection (no., % with condition)	No. (%)			
	Not hospitalized	Hospitalized, non-ICU	ICU admission	Hospitalization status unknown
Total with case report form (N = 74,439)	12,217	5,285	1,069	55,868
Missing or unknown status for all conditions (67,277)	7,074	4,248	612	55,343
<b>Total with completed information (7,162)</b>	<b>5,143</b>	<b>1,037</b>	<b>457</b>	<b>525</b>
One or more conditions (2,692, 37.6%)	1,388 (27)	732 (71)	358 (78)	214 (41)
Diabetes mellitus (784, 10.9%)	331 (6)	251 (24)	148 (32)	54 (10)
Chronic lung disease* (656, 9.2%)	363 (7)	152 (15)	94 (21)	47 (9)
Cardiovascular disease (647, 9.0%)	239 (5)	242 (23)	132 (29)	34 (6)
Immunocompromised condition (264, 3.7%)	141 (3)	63 (6)	41 (9)	19 (4)
Chronic renal disease (213, 3.0%)	51 (1)	95 (9)	56 (12)	11 (2)
Pregnancy (143, 2.0%)	72 (1)	31 (3)	4 (1)	36 (7)
Neurologic disorder, neurodevelopmental, intellectual disability (52, 0.7%) <sup>†</sup>	17 (0.3)	25 (2)	7 (2)	3 (1)
Chronic liver disease (41, 0.6%)	24 (1)	9 (1)	7 (2)	1 (0.2)
Other chronic disease (1,182, 16.5%) <sup>§</sup>	583 (11)	359 (35)	170 (37)	70 (13)
Former smoker (165, 2.3%)	80 (2)	45 (4)	33 (7)	7 (1)
Current smoker (96, 1.3%)	61 (1)	22 (2)	5 (1)	8 (2)
None of the above conditions <sup>¶</sup> (4,470, 62.4%)	3,755 (73)	305 (29)	99 (22)	311 (59)

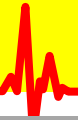
**Abbreviation:** ICU = intensive care unit.

\* Includes any of the following: asthma, chronic obstructive pulmonary disease, and emphysema.

<sup>†</sup> For neurologic disorder, neurodevelopmental, and intellectual disability, the following information was specified: dementia, memory loss, or Alzheimer's disease (17); seizure disorder (5); Parkinson's disease (4); migraine/headache (4); stroke (3); autism (2); aneurysm (2); multiple sclerosis (2); neuropathy (2); hereditary spastic paraplegia (1); myasthenia gravis (1); intracranial hemorrhage (1); and altered mental status (1).

<sup>§</sup> For other chronic disease, the following information was specified: hypertension (113); thyroid disease (37); gastrointestinal disorder (32); hyperlipidemia (29); cancer or history of cancer (29); rheumatologic disorder (19); hematologic disorder (17); obesity (17); arthritis, nonrheumatoid, including not otherwise specified (16);

# Un exemple récent – le statut tabagique est-il lié à la sévérité de la COVID-19?



Effectifs observés (nombre de cas)			
	Severe COVID	Mild COVID	Effectif Total
Smokers / past smokers	105	141	<b>246</b>
Non-smokers	1389	5002	<b>6391</b>
<b>% total</b>	<b>22.5%</b>	<b>77.5%</b>	



% observés			
	Severe COVID	Mild COVID	Total
Smokers / past smokers	43%	57%	<b>100%</b>
Non-smokers	22%	78%	<b>100%</b>

Effectifs théoriques que l'on devrait observer s'il n'y avait pas de relation (H0)			
	Severe COVID	Mild COVID	Effectif Total
Smokers / past smokers	55	191	<b>246</b>
Non-smokers	1439	4952	<b>6391</b>
<b>% total</b>	<b>22.5%</b>	<b>77.5%</b>	



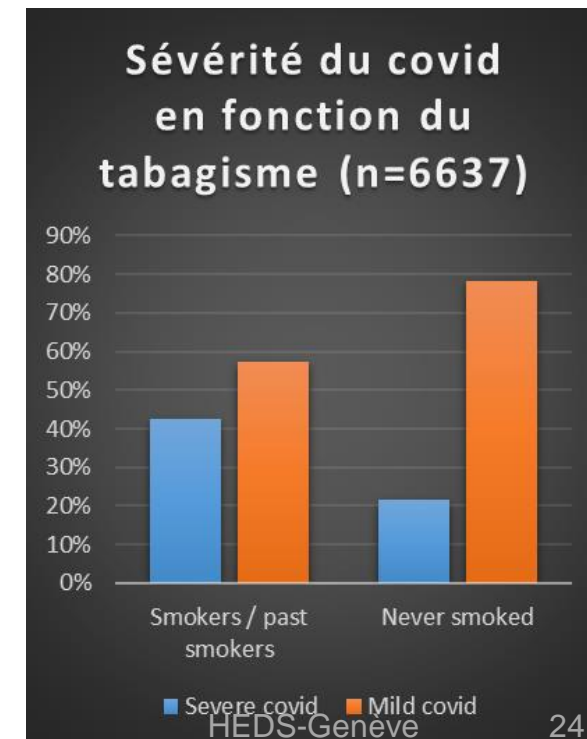
% théoriques s'il n'y avait pas de relation (H0)			
	Severe COVID	Mild COVID	Total
Smokers / past smokers	22.5%	77.5%	<b>100%</b>
Non-smokers	22.5%	77.5%	<b>100%</b>



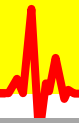
# Un exemple récent – le statut tabagique est-il lié à la sévérité de la COVID-19?

- Voici ces mêmes données illustrées par un graphique en barres
- Visuellement, ce pattern ressemble à une interaction. Le test du Chi-2 permet de tester statistiquement l'hypothèse d'une interaction (dépendance) entre ces deux facteurs catégoriels

	% observés		
	Severe COVID	Mild COVID	Total
Exposed	43%	57%	100%
Non-exposed	22%	78%	100%



# Le test du Chi2 se calcule en comparant les effectifs observés et théoriques



Effectifs observés (nombre de cas)			
	Severe COVID	Mild COVID	Effectif Total
Smokers / past smokers	105	141	246
Non-smokers	1389	5002	6391
% Total	22.5%	77.5%	

Effectifs théoriques que l'on devrait observer s'il n'y avait aucune relation (H0)			
	Severe COVID	Mild COVID	Effectif Total
Smokers / past smokers	55	191	246
Non-smokers	1439	4952	6391
% Total	22.5%	77.5%	

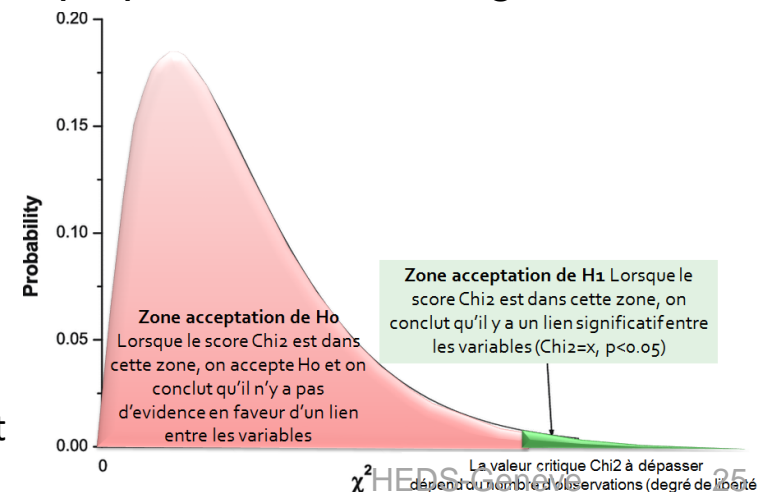
$$\text{Chi-2} = \sum \frac{(\text{obs} - \text{théo})^2}{\text{théo}}$$

**théo** : effectifs théoriques calculés

**obs** : effectifs observés dans l'échantillon

- $\chi^2 = (105-55)^2/55 + (141-191)^2/191 + \dots$
- $\chi^2 = 59.6$
- Seuil  $\chi^2$  à dépasser pour déclarer le test significatif ( $p < .05$ ) = 3.84
- H0 rejetée car le  $\chi^2$  calculé est > au seuil  $\chi^2$  à dépasser
- p-value calculée = 2.1e-14
- Conclusion: Le test est significatif ( $\chi^2 = 58.4, p < 0.001$ ), on conclut à un lien significatif entre le tabagisme et la sévérité du covid19

Une valeur de  $\chi^2$  très proche de 0 implique un test non-significatif



# Tableau de contingence et risque relatif (RR)

- Le risque relatif (RR), souvent utilisée en épidémiologie, mesure **la taille d'effet** du risque de survenue d'un évènement (maladie) chez des personnes exposées ou non à un facteur donné.
- C'est une mesure de la force d'association entre deux variables catégorielles qui peut venir compléter un test du  $\chi^2$

$$RR = \frac{A / (A+B)}{C / (C + D)}$$

- RR est une taille d'effet comprise entre  $]0, +\infty]$ 
  - Si  $RR = 1$ , l'exposition au facteur de risque n'est pas lié à l'outcome clinique (et le test de  $\chi^2$  n'est pas significatif)
  - Si  $RR > 1$  (facteur de risque)** l'exposition au facteur **est lié à une probabilité plus importante** de survenue de l'évènement dans le groupe exposé versus non-exposé, par un facteur RR
  - Si  $RR < 1$  (facteur protecteur)**, l'exposition au facteur **est lié à une probabilité moins importante** de survenue de l'évènement dans le groupe exposé versus non-exposé, par un facteur RR
- Dans notre exemple, le tabagisme est lié à un plus grand risque de développer une forme grave de la covid-19 d'un facteur **RR=1.96**

Effectifs observés (nombre de cas)		
	Severe COVID	Mild COVID
(past)Smokers	(A) 105	(B) 141
Non-smokers	(C) 1389	(D) 5002

# Tableau de contingence et le rapport des chances (odds ratio ou OR)

- Le rapport des chances (ou rapport des cotes, ou odds ratio (OR) en Anglais), est une autre mesure similaire de l'association (interaction) entre le facteur d'exposition et l'outcome clinique considéré.
- L'OR est le rapport entre la chance de survenue de la maladie dans le groupe exposé, et la chance de survenue de la maladie dans le groupe non-exposé
- L'OR est la mesure d'association utilisée dans la régression logistique
- Moins intuitif à comprendre que le RR

$$\text{OR} = \frac{A/B}{C/D} = \frac{\text{A chances contre B de survenue de la maladie dans le groupe exposé}}{\text{C chances contre D de survenue de la maladie dans le groupe non-exposé}}$$

- L'OR est une taille d'effet comprise entre ]0, +∞]
  - Si OR = 1, il n'y a pas d'association entre les deux facteurs
  - Si OR > 1 (**facteur de risque**) le rapport de chances est OR fois **plus grand** chez le groupe exposé que non-exposé
  - Si OR < 1 (**facteur protecteur**) le rapport de chances est OR fois **moins grand** chez le groupe exposé que non-exposé
- Dans notre exemple, l'OR = 2.68 IC95% [2.07, 3.47]
- Lorsque l'**intervalle de confiance** d'un rapport de chances comprend la valeur 1, on sait d'avance que le test statistique ne pourra rejeter l'hypothèse nulle (pas d'évidence en faveur d'un lien).



# Vidéo explicative – Taille d'effet RR et OR

Ep 11 : Taille d'effet. RR, OR, NNT ...

RISQUE  $\alpha$



# EP 11 : TAILLE D'EFFET

YouTube

0:00 / 17:22



# Plan et objectifs pédagogiques du cours

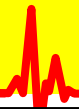


L'objectif pédagogique général du cours est de donner **une vue d'ensemble** de la notion de test statistique et des hypothèses qu'ils permettent de tester.

Nb: Notez que des notions comme les conditions d'applicabilité du test, le calcul et la distribution de la statistique, ou encore les notions de tests paramétriques et non-paramétriques sont des notions importantes non abordées dans ce cours

- Pouvoir expliquer la logique globale d'un test d'hypothèse, sans mathématiques
- Pouvoir expliquer le but d'un test d'association (encore appelé test de corrélation) et le but d'un test de différence de moyennes
  1. Association: Comprendre la logique du test du  $\chi^2$  (**Chi2**)
  2. Association: Comprendre la logique du test de **corrélation linéaire**
  3. Association: Comprendre la logique du test de **régression linéaire**
  4. Différence de moyennes : Comprendre la logique du **test de Student** et de **l'ANOVA**
- Pouvoir expliquer la notion de taille d'effet et son lien par rapport à la pertinence clinique d'une intervention

# Tests de corrélation et de régression

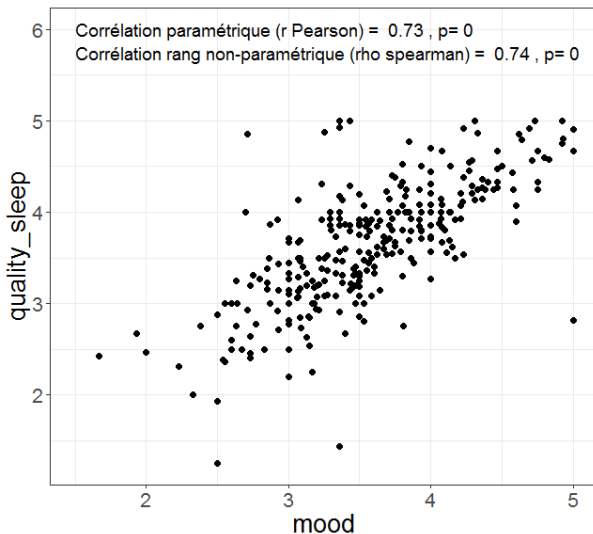


Type de Variable	Tests d'association (paramètre du test)		Tests de différence de moyennes (paramètre du test) NA = non-appariés; A = appariés	
	Lien ou corrélation entre 2 variables	Régression entre 2 variables	Entre 2 moyennes	Entre +2 moyennes
Catégorielle (sexe: ♀:10, ♂:15)	Test Chi-carré ( $\chi^2$ ) Test Phi ( $\phi$ ) Test exact de Fisher (F)	Régression logistique  (OR =>odds ratio; RR=>risque relatif)	NA: Chi-carré ( $\chi^2$ ) A: Chi-carré de McNemar ( $\chi^2$ )	NA: Chi-carré ( $\chi^2$ ) A: Cochran
Discrète (note: [1,2,3,4,5,6])	Rho Spearman ( $\rho$ ) Tau de Kendall ( $\tau$ )	Régression linéaire (R ou R <sup>2</sup> et b ou $\beta$ )	NA et A: Test de Mann-Whitney-Wilcoxon (U)	NA: Kruskal-Wallis (H) A: Friedman (Q)
Continue (poids: 70,3 kg)	Corrélation de Pearson (r)		NA et A: Test T de Student (t)	NA: ANOVA A: ANOVA à mesures répétées (F)

# Tester l'hypothèse d'une liaison entre deux variables

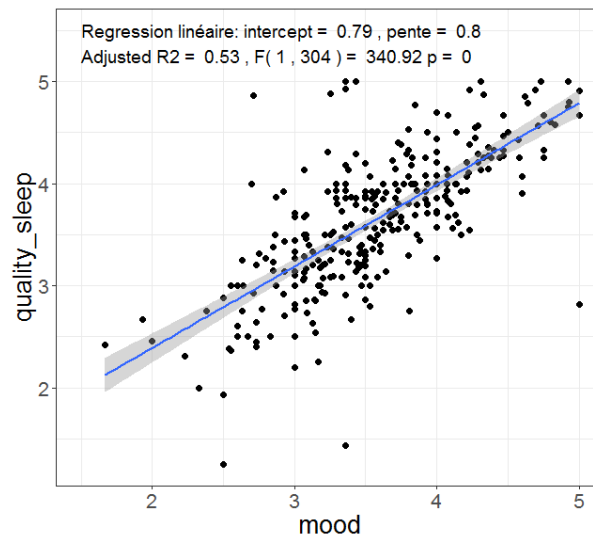
- 306 patients ont estimé la qualité de leur **sommeil** (1-8) et de leur **humeur** (1-5) durant le mois passé
- Ci-dessous, l'hypothèse d'une association entre sommeil et humeur a été testée de trois façons différentes sur les mêmes données.
- Conclusion: les trois tests donnent des résultats concordants, indiquant une association positive entre l'humeur et la qualité du sommeil

## Corrélation (Pearson r) VI et VD continues



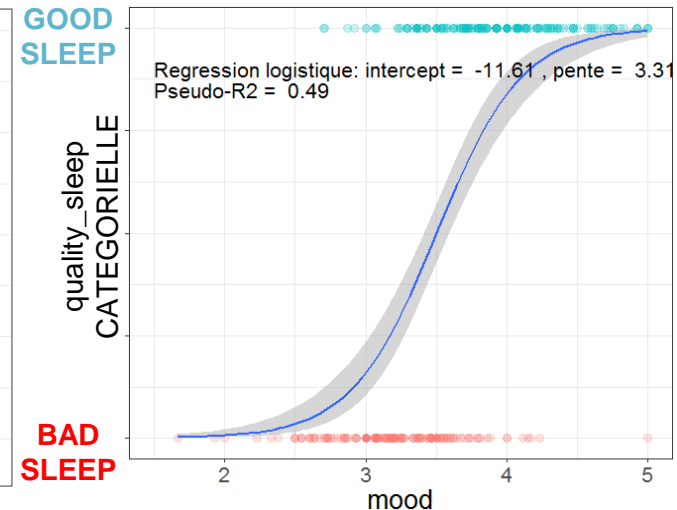
Résultat: je sais que ces deux variables sont associées avec une force d'association qui vaut r

## Régression linéaire VI et VD continues



Résultat: je sais que ces deux variables sont associées et je dispose d'une équation linéaire ou logistique qui me permet de prédire la qualité du sommeil (VD) à partir de l'humeur (VI)

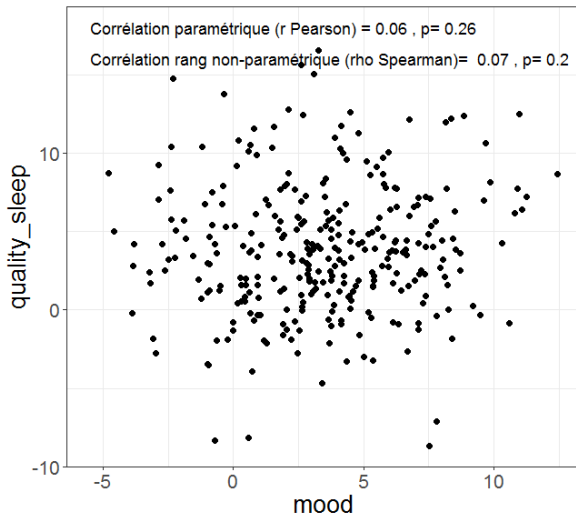
## Régression logistique VI catégorielle, VD continue



# Tester l'hypothèse d'une liaison entre deux variables

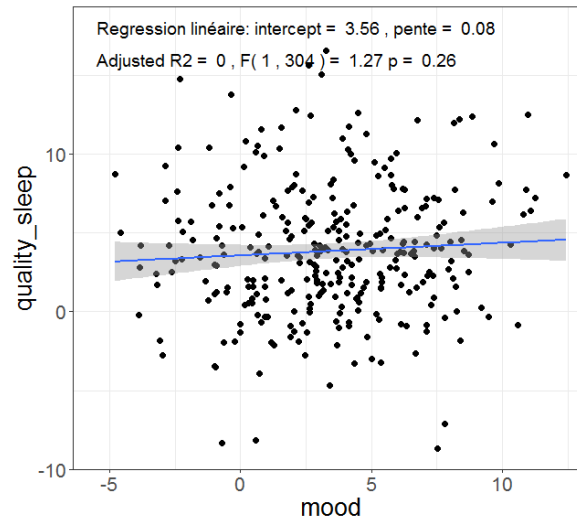
- Si ces deux mêmes variables n'étaient pas associées - ce que nous pouvons faire en multipliant chaque observation par un chiffre aléatoirement tiré entre 0 et 1 - voici ce que nous pourrions observer pour chacun des tests
- Le nuage de point serait sphérique => pas de lien
- Les régressions auraient une pente nulle/plate => pas de lien

## Corrélation (Pearson r) VI et VD continues



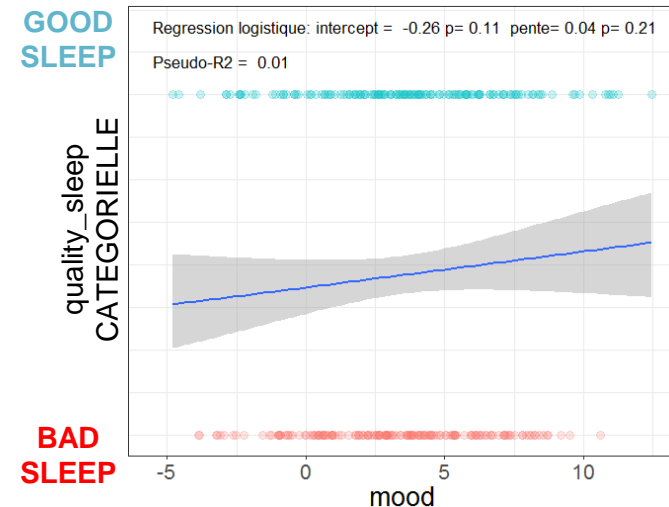
Les deux variables ne sont pas corrélées

## Régression linéaire VI et VD continues



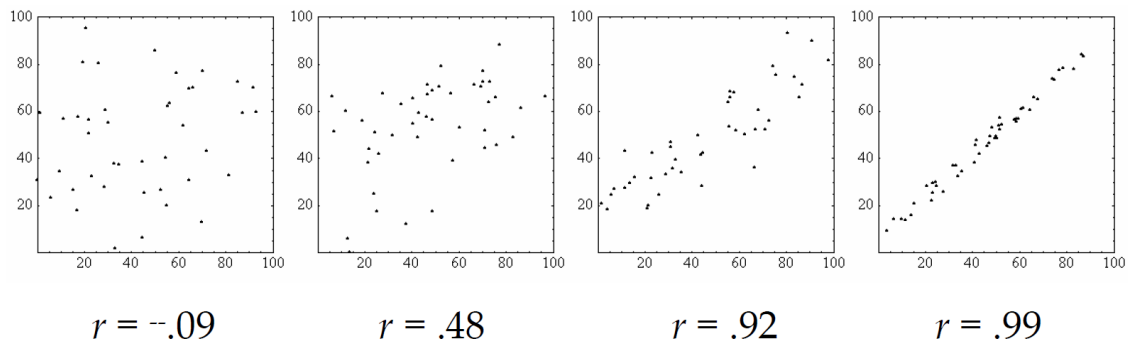
Les pentes des régressions ne sont pas significativement différentes de 0. Il n'y a donc pas d'association significative entre

## Régression logistique VI catégorielle, VD continue

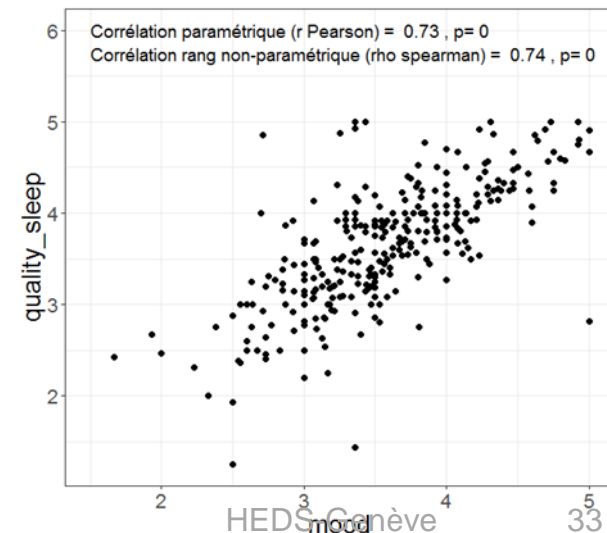


# Corrélation de Pearson

- Une corrélation entre deux variables a pour objectif d'estimer la force et la direction de l'association entre deux variables via un coefficient de corrélation ( $r$  ou rho)
- Le test de corrélation vise à tester si ce coefficient est significativement différent de 0, ce qui dépendra de sa force et du nombre d'observation
- Le coefficient de corrélation de Pearson ( $r$ ) est une mesure du degré et de la direction d'association entre deux variables. Il est compris entre  $[-1; 1]$ 
  - Si  $r$  tend vers  $-1$ : corrélation négative
  - Si  $r$  tend vers  $+1$ : corrélation positive
  - Si  $r$  tend vers  $0$ : corrélation nulle

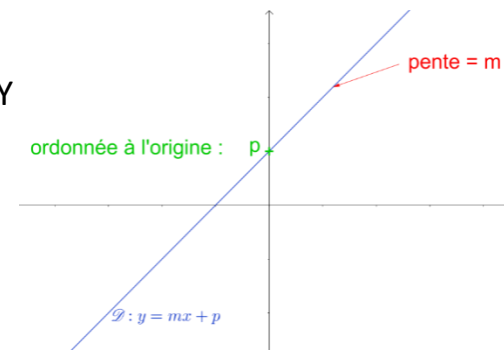
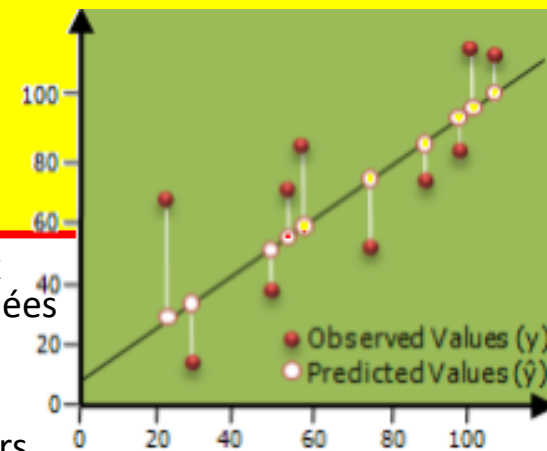


- Hypothèses du test de corrélation
  - $H_0$ : Le coefficient de corrélation  $r$  ne diffère pas de 0
  - $H_1$ : Le coefficient de corrélation diffère de 0



# La régression linéaire simple

- Une régression linéaire vise à estimer l'équation de la droite qui colle le mieux aux données, c'est-à-dire qui passe le plus près des points d'un nuage de données
- Cette droite est un modèle mathématique simple qui permet de résumer les données
- L'équation de cette droite est la droite qui minimise la distance entre les valeurs observées (en rouge), et les valeurs prédites par la droite (en blanc).
- L'équation d'une droite se définit à l'aide de deux paramètres :
  - **L'ordonnée à l'origine (intercept en Anglais)** de la droite qui désigne la valeur de Y lorsque X = 0
  - **La pente de la droite** (appelé aussi coefficient directeur) qui désigne son inclinaison. Si la pente tend vers 0, la droite est « plate » et le modèle ne prédit rien => il n'existe pas de relation linéaire / d'association entre les deux variables

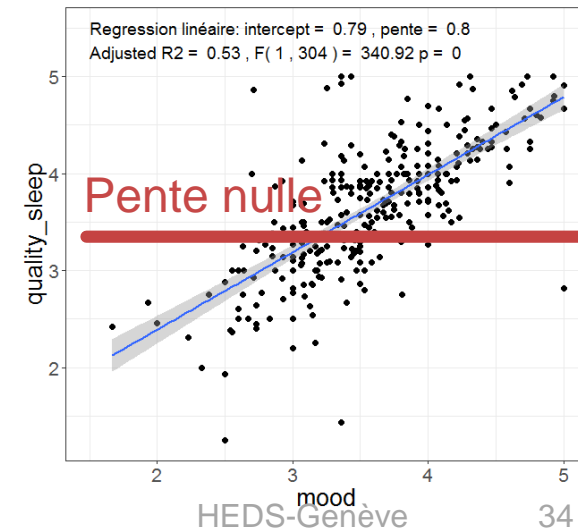


$$Y = \text{pente} * X + \text{ordonnée\_à\_l'origine} + \text{erreur}$$

Ce qui est connu / observé

Paramètres que l'on cherche à estimer

- Hypothèses pour le modèle linéaire dans son ensemble (F stat.)
  - H0: Le modèle estimé ne prédit pas mieux les données **que le modèle nul**
  - H1: Le modèle estimé prédit mieux les données **que le modèle nul**
- Hypothèses pour chacun des paramètres (t student stat.)
  - H0: Les paramètres de pente (resp. d'ordonnée\_origine) ne diffèrent pas de 0
  - H1: Les paramètres de pente (resp. d'ordonnée\_origine) diffèrent de 0





# Plan et objectifs pédagogiques du cours

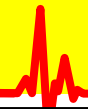


L'objectif pédagogique général du cours est de donner **une vue d'ensemble** de la notion de test statistique et des hypothèses qu'ils permettent de tester.

Nb: Notez que des notions comme les conditions d'applicabilité du test, le calcul et la distribution de la statistique, ou encore les notions de tests paramétriques et non-paramétriques sont des notions importantes non abordées dans ce cours

- Pouvoir expliquer la logique globale d'un test d'hypothèse, sans mathématiques
- Pouvoir expliquer le but d'un test d'association (encore appelé test de corrélation) et le but d'un test de différence de moyennes
  1. Association: Comprendre la logique du test du  $\chi^2$  (**Chi2**)
  2. Association: Comprendre la logique du test de **corrélation linéaire**
  3. Association: Comprendre la logique du test de **régression linéaire**
  4. Différence de moyennes : Comprendre la logique du **test de Student** et de **l'ANOVA**
- Pouvoir expliquer la notion de taille d'effet et son lien par rapport à la pertinence clinique d'une intervention

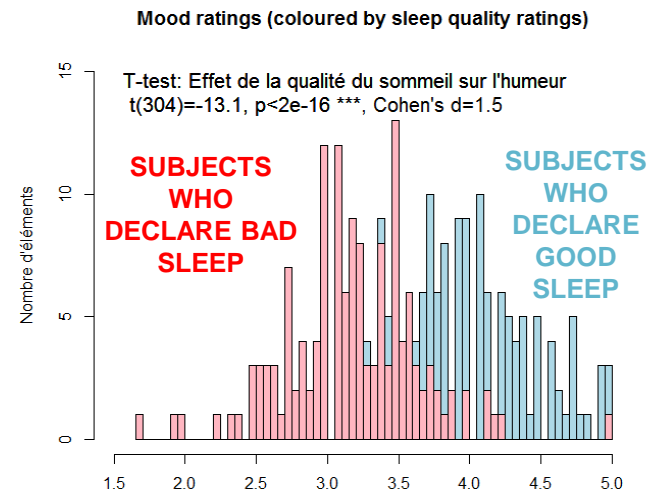
# Tests de comparaison de moyennes



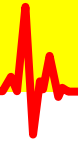
Type de Variable	Tests d'association (paramètre du test)		Tests de différence de moyennes (paramètre du test) NA = non-appariés; A = appariés	
	Lien ou <b>corrélation</b> entre 2 variables	Régression entre 2 variables	Entre 2 moyennes	Entre +2 moyennes
Catégorielle (sexe: ♀:10, ♂:15)	Test Chi-carré ( $\chi^2$ ) Test Phi ( $\phi$ ) Test exact de Fisher ( <b>F</b> )	Régression logistique  ( <b>OR</b> => odds ratio; <b>RR</b> => risque relatif)	NA: Chi-carré ( $\chi^2$ )  A: Chi-carré de McNemar ( $\chi^2$ )	NA: Chi-carré ( $\chi^2$ )  A: Cochran ( <b>Q</b> )
Discrète (note: [1,2,3,4,5,6])	Rho Spearman ( <b><math>\rho</math></b> ) Tau de Kendall ( <b><math>\tau</math></b> )	Régression linéaire ( <b>R</b> ou <b>R<sup>2</sup></b> et <b>b</b> ou <b><math>\beta</math></b> )	NA et A: Test de Mann-Whitney-Wilcoxon ( <b>U</b> )	NA: Kruskal-Wallis ( <b>H</b> )  A: Friedman ( <b>Q</b> )
Continue (poids: 70,3 kg)	<b>Corrélation de Pearson (<b>r</b>)</b>		NA et A: Test T de Student ( <b>t</b> )	NA: ANOVA  A: ANOVA à mesures répétées ( <b>F</b> )

# Tester l'hypothèse d'une différence de moyennes entre deux échantillons (VI catégorielle, VD continue)

- Le test-t de Student a pour but de tester l'hypothèse d'une différence entre les moyennes de deux échantillons
  - Si les échantillons sont non-appariés (mesures issues de groupes distincts, unpaired en Anglais): on utilise le t-test non-apparié
  - Si les échantillons sont appariés (mesures issues du même groupe, paired en Anglais): on utilise le t-test apparié
- Ex: On applique le test-t de Student pour tester l'hypothèse d'une différence entre les moyennes d'humeurs mesurées dans ces deux échantillons
  - H0: L'humeur ne diffère pas en fonction de la qualité du sommeil
  - H1: L'humeur diffère en fonction de la qualité du sommeil
- Résultats concordants avec les tests de corrélation 😊



# La statistique de t du test de Student mesure une différence de moyenne normalisée par la variance des échantillons



$$t = \frac{\text{signal}}{\text{noise}} = \frac{\text{difference in means}}{\text{sample variability}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$\bar{x}_1, \bar{x}_2$  = sample means

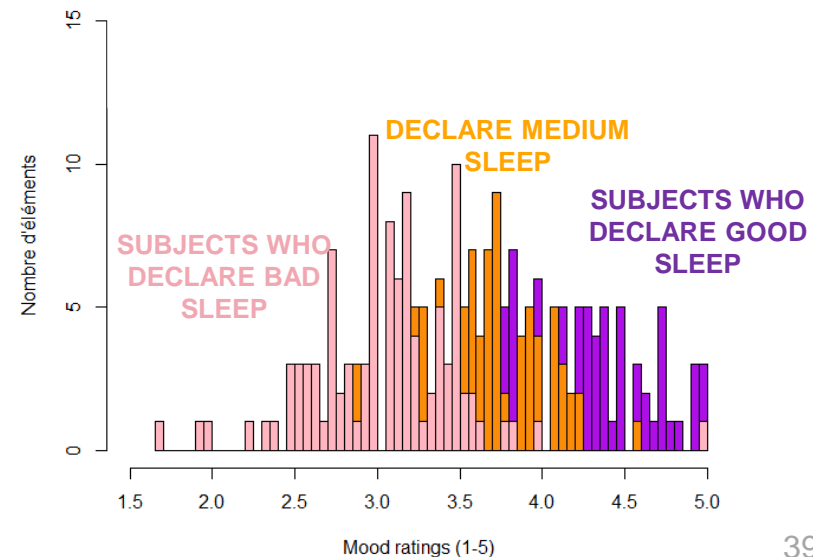
$s_1^2, s_2^2$  = sample variances

$n_1, n_2$  = sample sizes

# L'ANOVA pour tester l'hypothèse d'une différence de moyennes entre au moins trois échantillons



- Echantillons non-appariés (groupes distincts): on utilise l'ANOVA
- Echantillons appariés (issus du même groupe): ANOVA à mesures répétées
  
- On forme trois groupes d'individus selon la qualité de leur sommeil (mauvais, intermédiaire, bon)
- On pose la question de savoir si l'humeur moyenne de ces trois groupes diffère l'une de l'autre
  - H0: L'humeur moyenne entre les 3 groupes n'est pas significativement différente
  - H1: Au moins une moyenne diffère des autres
  
- L'ANOVA est significative – résultats concordants avec les autres tests 😊
- $\eta^2$  représente la taille d'effet dans les ANOVA – elle est importante



# Plan et objectifs pédagogiques du cours



L'objectif pédagogique général du cours est de donner **une vue d'ensemble** de la notion de test statistique et des hypothèses qu'ils permettent de tester.

Nb: Notez que des notions comme les conditions d'applicabilité du test, le calcul et la distribution de la statistique, ou encore les notions de tests paramétriques et non-paramétriques sont des notions importantes non abordées dans ce cours

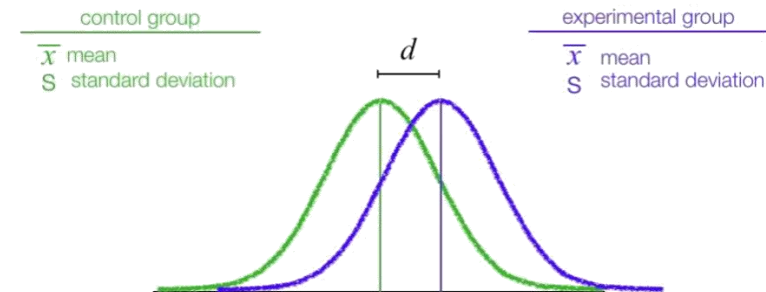
- Pouvoir expliquer la logique globale d'un test d'hypothèse, sans mathématiques
- Pouvoir expliquer le but d'un test d'association (encore appelé test de corrélation) et le but d'un test de différence de moyennes
  1. Association: Comprendre la logique du test du  $\chi^2$  (**Chi2**)
  2. Association: Comprendre la logique du test de **corrélation linéaire**
  3. Association: Comprendre la logique du test de **régression linéaire**
  4. Différence de moyennes : Comprendre la logique du **test de Student** et de **l'ANOVA**
- Pouvoir expliquer la notion de taille d'effet et son lien par rapport à la pertinence clinique d'une intervention



# La notion de taille d'effet (TE) quantifie la magnitude d'un effet, ce qui permet d'évaluer la pertinence clinique d'une intervention

- La **taille d'effet (TE)** désigne la **force d'un phénomène**, c'est-à-dire la force d'association entre deux variables ou la grandeur de la différence entre deux moyennes à l'aide d'un indicateur standardisé. Il existe différentes tailles d'effets
- La significativité statistique (p) ne donne pas d'information quant à la taille de l'effet considéré (ex: la p value est fonction de la taille d'effet et de l'effectif)

- **TE d'un lien entre variables catégorielles:** RR ou OR
- **TE d'une corrélation:** r de Pearson et rho de Spearman
- **TE d'un facteur de régression:** R2 ou f2
  
- **TE différence entre deux moyennes:** d de Cohen
- **TE d'un facteur d'ANOVA:** Eta<sup>2</sup>



$$d = \frac{|m_1 - m_2|}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

- **Pourquoi est-ce important?**
  - Parce qu'elle quantifie la force, et donc la pertinence clinique d'une intervention en fonction du gain escompté (et non en fonction de la significativité, qui ne dit rien sur la taille de l'effet). Introduire une nouvelle intervention prend du temps, des ressources etc. il est donc important de pouvoir estimer le bénéfice réel pour le patient-e
  - Les tailles d'effets permettent de **conduire des méta-analyses** pour synthétiser la taille d'effet moyenne d'une famille d'intervention
  - Les tailles d'effets permettent enfin de **conduire des analyses de puissance** avant de débiter une expérience. Ces analyses visent à déterminer la taille de l'échantillon qu'il sera nécessaire de recruter pour être sûr de détecter un effet avec une taille d'effet donnée et une puissance donnée (le risque beta de déclarer un faux négatif, souvent fixé à 80%). Une étape obligatoire pour justifier un financement de projet de recherche

# Le d de Cohen – Une taille d'effet utilisée pour quantifier la différence standardisée entre deux moyennes

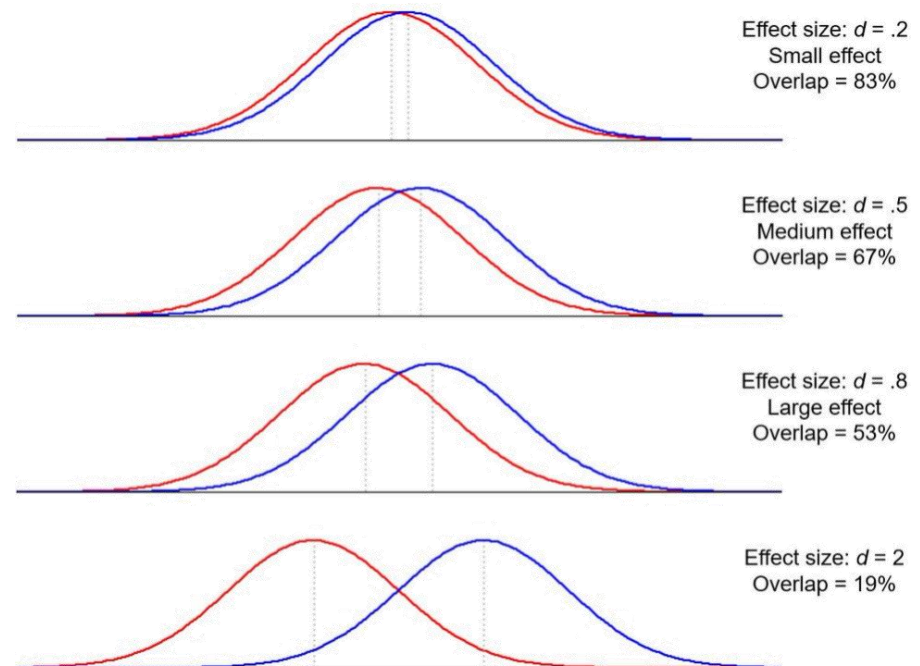
- Le *d de Cohen* quantifie une différence de moyennes en unité de déviation standard
- C'est la raison pour laquelle le *d de Cohen* est appelé la différence des moyennes standardisées
- Si  $d = x$ , on dit que les deux moyennes sont séparées de  $x$  déviation standard

Echantillons indépendants  
(2 groupes distincts)

$$d = \frac{|m_1 - m_2|}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

différence de moyennes  
-----  
déviation standard moyenne

m: moyenne  
s: variance



# Ordre de grandeur des principales tailles d'effets



- Pour information, ce tableau permet d'interpréter la magnitude des différentes tailles d'effets

<i>Test statistique</i>	<i>Estimateur de taille d'effet</i>	<i>Petite</i>	<i>Moyenne</i>	<i>Grande</i>
Khi2	<i>V de Cramer [0, 1] ou Phi [-1, 1]</i>	0.1	0.3	0.5
Corrélation	<i>r [-1, 1]</i>	0.1	0.3	0.5
Différences entre deux moyennes, t-tests	<i>d de cohen [-∞, +∞]</i>	0.2	0.5	0.8
ANOVA (éta-carré ou éta-carré partiel, ou oméga-carré)	$\eta^2$ ]0, 1], $\eta^2_p$ ]0, 1] ou $\omega^2$ ]0, 1]	0.01	0.06	0.14
Regression, ANOVA	$f^2$ ]0, 1]	0.02	0.15	0.35
Source: Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159.				
Convertisseur de tailles d'effets: <a href="http://www.stat-help.com/spreadsheets/Converting%20effect%20sizes%202012-06-19.xls">http://www.stat-help.com/spreadsheets/Converting%20effect%20sizes%202012-06-19.xls</a>				
Plus de tailles d'effets: <a href="https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize">https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize</a>				

# Exemple tiré de votre exercice 3.2 – Sur quel facteur de stress choisiriez-vous d’agir en premier ?

- Ce tableau présente 5 variables de stress mesurées chez des soignants canadiens exerçant en soin palliatifs (SP) soit en milieu hospitalier (CH) soit dans les EMS (CSSS)
- On vous invite à vous prononcer sur des actions visant à réduire le stress en milieu hospitalier (CH). Sur la base des chiffres suivants, quel est le facteur de stress sur lequel vous proposeriez d’agir en premier?

**Tableau 2** Description du stress au travail et analyse de variance pour comparer le stress des infirmières selon le milieu de pratique (CSSS et CH) dans le cadre de l’étude 1 ( $n = 209$ ) et de l’étude 2 ( $n = 751$ )

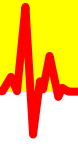
Variables de stress	Étendue	Total	CH	CSSS	$t^a$	$p$	$d$ de Cohen <sup>b</sup>
		$m$ (ET)	$m$ (ET)	$m$ (ET)			
<b>Demands</b>							
Demands psychologiques	9–36	28,21 (3,78)	29,28 (3,68)	26,80 (3,43)	9,36*	0,000	0,66
Efforts	6–24	18,37 (3,11)	18,96 (3,11)	17,57 (2,94)	6,19*	0,000	0,45
<b>Ressources</b>							
Latitude décisionnelle	24–96	73,17 (8,55)	72,60 (8,28)	73,68 (8,85)	–1,70	0,089	0,13
Récompenses	11–44	31,94 (4,79)	32,32 (4,71)	31,43 (4,84)	2,53+	0,012	0,19
Satisfaction au travail	1–7	4,96 (1,10)	4,94 (1,04)	4,99 (1,12)	–0,611	0,541	0,05

CH : centre hospitalier ; CSSS : centres de santé et de services sociaux ; ET : écart-type ; \* :  $p < 0,01$ .

<sup>a</sup> Un seuil significatif de 0,01 a été retenu afin de contrôler l’erreur de type I (approche de Bonferonni).

<sup>b</sup> Taille d’effet.

# Les méthodologistes en science infirmière argumentent en faveur de toujours rapporter la taille d'effet dans un test statistique



> J Nurs Educ. 2017 Nov 1;56(11):645-647. doi: 10.3928/01484834-20171020-02.

## The Purpose and Power of Reporting Effect Sizes in Nursing Education Research

Darrell Spurlock Jr

PMID: 29091232 DOI: 10.3928/01484834-20171020-02

### Abstract

Like many other fields and subfields within the social sciences, education, and medicine, nursing education research has a long history of giving p-values associated with common statistical tests a position of primary importance in interpreting study findings. Global, transdisciplinary efforts are underway to diminish the role p-values play in making judgements about the significance of study findings by changing the threshold p-value used to pronounce statistical significance, through the use of Bayesian statistics, and **through efforts to report effect sizes alongside p-values. In this month's Methodology Corner installment, the focus is on effect sizes and their role in enhancing the value and utility of nursing education research studies.** [J Nurs Educ. 2017;56(11):645-647].

*À l'instar de nombreux autres domaines et sous-domaines des sciences sociales, de l'éducation et de la médecine, la formation à la recherche en soins infirmiers accorde depuis longtemps aux valeurs p des tests statistiques une place de premier plan dans l'interprétation des résultats des études. **Des efforts mondiaux et transdisciplinaires sont en cours pour diminuer le rôle que jouent les valeurs p dans les jugements sur l'importance des résultats des études, en modifiant le seuil de la valeur p utilisée pour prononcer la signification statistique (ex:  $p < 0.01$  plutôt que  $p < 0.05$ ), en utilisant les statistiques bayésiennes et en s'efforçant de signaler les tailles d'effet en même temps que les valeurs p. Dans le numéro de ce mois-ci, l'accent est mis sur les tailles d'effet et leur rôle dans l'amélioration de la valeur et de l'utilité des études de recherche sur la formation en soins infirmiers.***

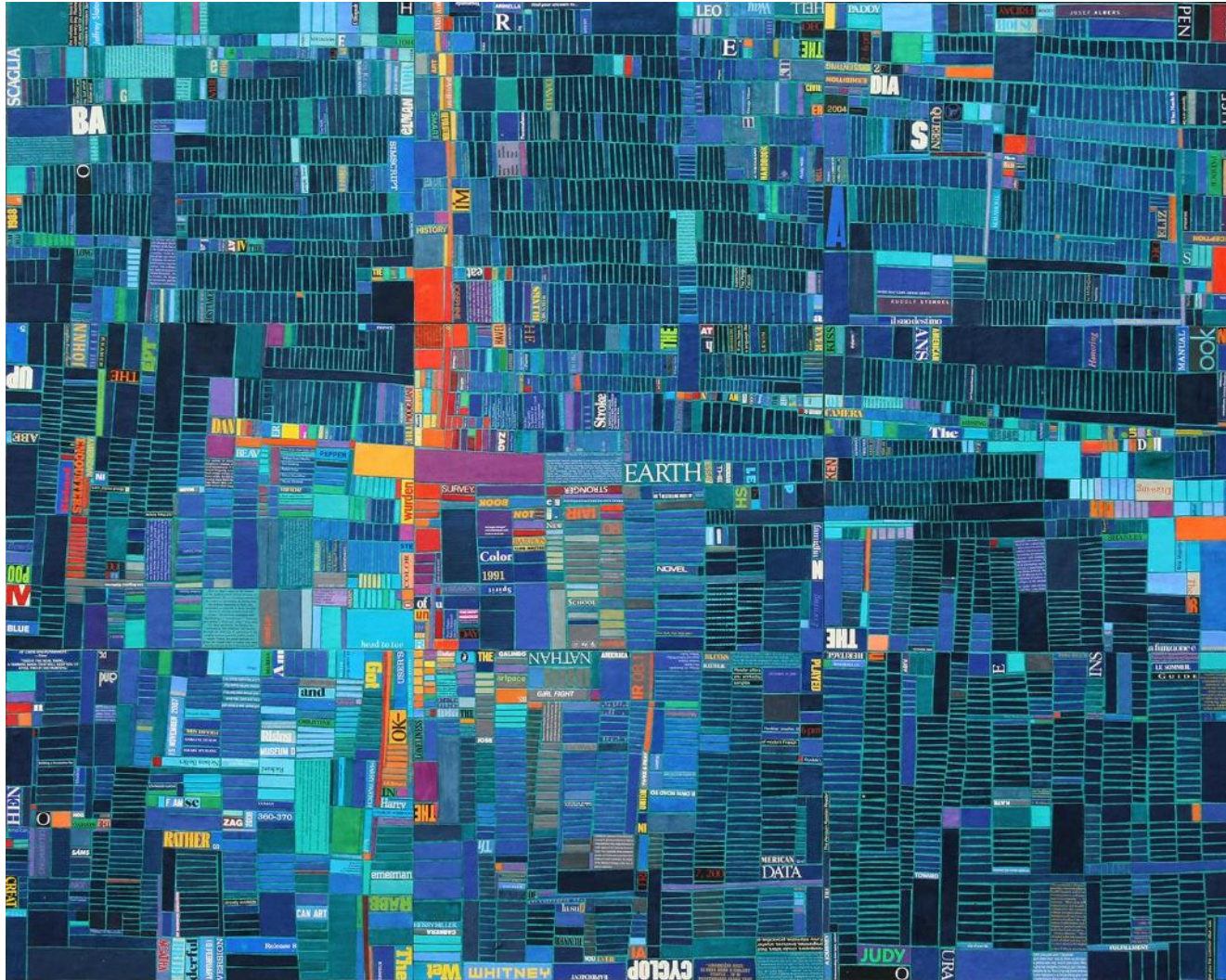
# Un dernier mot avant de finir



- L'objectif de ce cours était de vous faire comprendre la logique des principaux test statistiques et les inférences qu'elles permettent de tester
- Nous ne nous attendons pas à ce que vous reteniez tous les détails présentés aujourd'hui, ni à savoir appliquer ces tests concrètement, mais au moins à les lire et les comprendre
- **Nous vous encourageons à poser toutes les questions qui pourront vous aider et aider vos camarades à mieux comprendre les objectifs abordés dans ce cours**



# Merci de votre attention – et merci de poser toutes vos questions !



Imagined Time, Laura Frick

# Q1

- dans le premier quiz, la dernière question était au sujet de la fiabilité, quelle est la règle pour savoir si les pourcentages sont fiables ou pas ?

--> Je comprends tout à fait la réponse mais au final que signifie "plutôt faible", par rapport à quoi est-ce plutôt faible. Je m'explique un écart de 2% est faible comme peut aussi l'être un écart de 5%. A partir de quand est-ce que l'écart n'est plus faible ?

- Quiz: Concernant l'hypothèse que les étudiants en soins infirmiers de différentes années d'étude n'ont pas la même fréquence de stress, les différences de pourcentages entre les trois années sont-elles plutôt faibles ?
- Excellentes questions! Comment savoir si un effet est faible ou non? Par rapport à quoi?

Fréquence du stress déclaré en lien avec la formation	Non pas du tout	4,0%	2,8%	1,0%	2,6%	<0,001
	Non pas beaucoup	7,6%	5,0%	4,8%	5,7%	
	Oui parfois	62,9%	47,9%	43,2%	51,0%	
	Oui souvent ou tout le temps	25,6%	44,4%	50,9%	40,8%	
	N	450	505	495	1450	

**Table 1** Values of Effect Sizes and Their Interpretation

<i>Kind of Effect Size</i>	<i>Small</i>	<i>Medium</i>	<i>Large</i>
----------------------------	--------------	---------------	--------------

*Nous pourrions utiliser la taille d'effet du test de Khi2: V de Cramer [0, 1]. Cette taille d'effet désigne la force d'association entre les deux variables. Son interprétation est semblable au score r de la corrélation de Pearson*

.10                      .30                      .50

# Question2: De l'usage du mot clinique



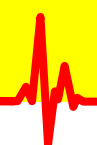
- Je suis pas sure d'avoir bien compris quand dire "statistiquement mais pas cliniquement meilleure" ou au contraire, "statistiquement et cliniquement meilleure".

- Je ne comprends pas cette question : peut-on dire que les consommateurs ASH sont cliniquement plus âgés que les non-consommateurs, mais que cette différence n'est pas significative ? Je ne comprends pas ce que veut dire « cliniquement » et du coup comment on peut savoir si la différence est significative ou non.

- Réponse: L'usage du mot clinique dans cet énoncé peut prêter à confusion. Le mot «cliniquement» dans ce contexte est utilisé comme synonyme de «numériquement» et ne revêt pas de qualité particulière => On observe que les consommateurs ASH sont numériquement plus âgés, mais cette différence n'est pas significative.

# Question3:

• Serait-ce possible de revoir le concept de "pratiquement ou cliniquement importante" ?



- La pertinence / l'importance clinique d'une intervention est un compromis entre la taille d'effet du bénéfice escompté pour le patient, le temps et les ressources nécessaires pour l'implémenter

**Table 1** Values of Effect Sizes and Their Interpretation

Kind of Effect Size		Small	Medium	Large
■ $\chi^2$	<i>V de Cramer [0, 1] ou Phi [-1, 1]</i>	.10	.30	.50
■ Correlation	<i>r [-1, 1]</i>	.10	.30	.50
■ Diff. entre deux moyennes	<i>d</i>	0.20	0.50	0.80
■ ANOVA (eta-carré partiel)	$\eta^2_p$	.01	.06	.14
■ Regression	$f^2$	.02	.15	.35

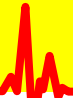
Source: Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. doi:10.1037/0033-2909.112.1.155

Convertisseur de tailles d'effets:

<http://www.stat-help.com/spreadsheets/Converting%20effect%20sizes%202012-06-19.xls>

## Quizz 3.1.2 Lire et interpréter une corrélation et un OR présentés dans un article sous forme de tableau (question 9 et 10)

## Quizz 3.1.3 Lire et interpréter un t-test présenté dans un article sous forme de tableau (questions 3 et 7)



Je vous écris au sujet du premier exercice de la partie "Test statistique et taille d'effet", le quizz ce nomme "lire et interpréter un Chi<sup>2</sup>" que vous avez ouvert lundi. Pour les questions suivantes :

5. L'hypothèse que les étudiants de 3ème année sont plus nombreux que ceux de 1ère et 2ème année à rapporter être stressés souvent ou tout le temps est-elle corroborée, si l'on prend comme seuil .05?

--> Ici les 3ième année obtiennent une valeur de 50,9% pour 495 élèves (qui est donc supérieur aux deux autres années avec 44,4% de 505 élèves et 25,6% de 450 élèves).  $P < 0,001$  donc  $< 0,05$ . J'ai donc répondu que effectivement l'hypothèse me semblait être corroborée. Mais ce n'est pas le cas et j'aimerais comprendre pourquoi ?

- Réponse: Non car même si la valeur p est inférieure à .05 elle ne teste pas cette hypothèse. L'hypothèse testée par le Khi<sup>2</sup> est celle que la fréquence du stress (en considérant toutes les catégories de réponse) diffère selon l'année d'étude



# → Quizz 3.1.4 lire et interpréter un intervalle de confiance et une taille d'effet présentés dans un article sous forme de tableau (questions 6 et 7)



## Questions TPE Test statistiques et taille d'effet

Je n'ai pas du tout compris la taille d'effet et l'intervalle de confiance, malgré le visionnage du PPT et de la vidéo Youtube en annexe, et mes résultats de l'exercice 3.1.4 le démontrent. Par exemple, je n'ai pas compris comment à partir d'un tableau de corrélation on peut savoir la taille d'effet ?

Pour la ligne « WAI lien », peut-on conclure que la corrélation entre alliance thérapeutique et pénibilité perçue des contrôles d'urine est forte ?

Oui  Non

Note de 0,00 sur 1,00

La réponse correcte est : Oui

Tableau 4  
Corrélations

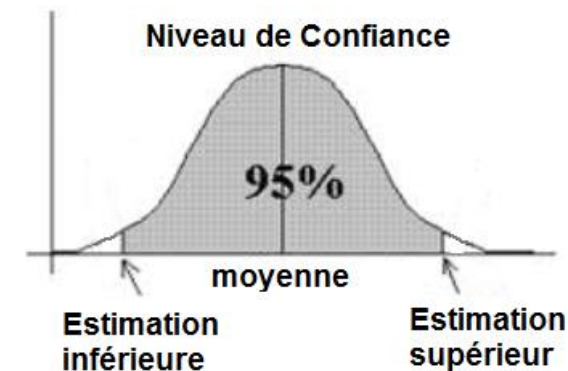
Variables	Corrélation Pearson	Significativité (valeur-p)	Corrélation de Spearman-Brown (rho)
WAI taux total/VAS	-0.67	< 0.001	-0.69
WAI but/VAS	-0.75	< 0.001	-0.80
WAI tache/VAS	-0.63	< 0.001	-0.64
WAI lien/VAS	-0.48	0.008	-0.51

- Réponse: Le r de Pearson est une taille d'effet



# La notion d'intervalle de confiance

- Lorsque l'on estime un paramètre sur un échantillon, l'objectif est souvent d'étendre ce résultat à l'ensemble de la population, que nous ne pouvons mesurer dans son ensemble.
- Au niveau de la population, l'estimation de ce paramètre implique nécessairement une marge d'erreur.
- Plus le nombre d'observations augmentent, plus l'estimation est précise et la marge d'erreur diminue.
- L'intervalle de confiance se calcule toujours en fonction d'un certain niveau de confiance, ici 95%, la valeur la plus fréquemment utilisée.
  - Ex: L'Age moyen des 20 premiers étudiants de l'amphoithéatre, tirés au hasard est de 22.4 ans IC95% [21.8 – 23]
- Cela veut dire que si l'on devait mesurer l'Age des 20 étudiants suivants tirés dans la même population, nous aurions 95% de chance de tomber sur une moyenne comprise entre [21.8 – 23]



- Une autre chose que je n'ai pas compris sont les premières questions qui reviennent à chaque exercice. Je ne comprends pas le sens de la question et ce qui est réellement recherché. J'ai eu de la difficulté pour tous les exercices qui contenaient ce type de question.

Facteurs associés à la poursuite de l'allaitement jusqu'à 6 mois chez les mères allaitantes dans une maternité parisienne

Caractéristiques	OR (IC95%)
Détermination à allaiter Non (n=44) Oui (n=37)	1 [référence] 3,99 (1,04-15,31)
Difficulté à téter à 48 h Non (n=51) Oui (n=30)	1 [référence] 0,20 (0,05-0,77)
Introduction du lait commercial avant 3 mois Non (n=46) Oui (n=35)	1 [référence] .06 (0,01-0,29)
Anxiété de la mère à 48h (n=81)	1,18 (1,06-1,32)
Sensibilité-Engagement à 48 h (n=81)	0,16 (0,03-0,84)

Tableau 3. Modèle de régression logistique des facteurs associés à la poursuite de l'allaitement maternel à six mois auprès de 81 dyades mères-enfants

Réponse: La question est difficile ne vous inquiétez pas

A quelles hypothèses de recherche ou alternatives pourrait se référer l'intervalle de confiance à 95% qui figure dans le tableau pour « Difficulté à téter à 48h » ? ATTENTION: nous cherchons toutes les hypothèses possibles qui auraient pu amener à faire les tests statistiques présentés ici, indépendamment qu'elles soient confirmées ou non.

Veuillez choisir au moins une réponse :

1. Il n'y a pas de différence de poursuite de l'allaitement à 6 mois en fonction de la difficulté à téter à 48h
2. Le nombre moyen de mois d'allaitement diffère entre les mères dont les enfants ont eu ou non des difficultés à téter à 48h
3. Les difficultés à téter de l'enfant à 48h influencent le fait que la mère poursuit ou non l'allaitement à 6 mois
4. Les mères dont l'enfant a eu une difficulté à téter à 48h sont plus nombreuses à avoir arrêté d'allaiter à 6 mois
5. Les mères dont l'enfant a eu une difficulté à téter à 48h sont plus nombreuses à avoir continué d'allaiter à 6 mois
6. Les difficultés à téter de l'enfant à 48h prédisent la détermination à allaiter (SP: QUESTION PIEGE ... IL MANQUE «A 6 MOIS»)
7. La poursuite ou non de l'allaitement maternel à 6 mois est prédite par la difficulté ou non à téter de l'enfant à 48h

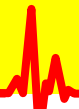
# Exercice 3.1.1: Lire et interpréter un Chi-carré présenté dans un article sous forme de tableau



Tableau 2  
Santé déclarée, stress et Comportements de santé selon l'année d'étude

		1° année d'étude	2° année d'étude	3° année d'étude	Total	p
Etat de santé global déclaré	Mauvaise	1,1%	3,2%	5,1%	3,2%	<0,01
	Correcte	28,7%	33,5%	35,4%	32,7%	
	Bonne	41,9%	42,2%	38,7%	40,9%	
	Très bonne	19,4%	14,9%	16,0%	16,7%	
	Excellente	8,9%	6,2%	4,9%	6,6%	
	N	449	502	494	1445	
Fréquence du stress déclaré en lien avec la formation	Non pas du tout	4,0%	2,8%	1,0%	2,6%	<0,001
	Non pas beaucoup	7,6%	5,0%	4,8%	5,7%	
	Oui parfois	62,9%	47,9%	43,2%	51,0%	
	Oui souvent ou tout le temps	25,6%	44,4%	50,9%	40,8%	
	N	450	505	495	1450	
Sport	Non	52,0%	56,8%	58,0%	55,7%	ns
	Oui	47,8%	43,0%	42,0%	44,1%	
	N	449	504	495	1448	
Tabac	Non	61,6%	62,0%	61,2%	61,6%	ns
	Oui, occasionnellement	8,9%	9,1%	9,7%	9,2%	
	Oui, tous les jours	29,3%	28,9%	28,9%	29,0%	
	N	449	505	494	1448	
Alcool	Non	38,7%	36,4%	35,8%	36,9%	ns
	Oui, occasionnellement	60,4%	62,4%	62,2%	61,7%	
	Oui, tous les jours	0,9%	1,2%	2,0%	1,4%	
	N	450	505	495	1450	

# Exercice 3.1.2 - Lire et interpréter une corrélation et un OR présentés dans un article sous forme de tableau



- A quelles hypothèses de recherche ou alternatives pourrait se référer la valeur p qui figure dans le tableau pour la ligne « WAI taux total » ? ATTENTION: nous cherchons toutes les hypothèses possibles qui auraient pu amener à faire les tests statistiques présentés ici, indépendamment qu'elles soient confirmées ou non. Veuillez choisir au moins une réponse :
- **1. Les patients en traitement de dépendances qui ont une meilleure alliance trouvent les contrôles d'urine moins pénibles**
- 2. Les patients en traitement de dépendances ont plus d'alliance que de pénibilité perçue des contrôles d'urine
- 3. Le pourcentage de patients en traitement de dépendances avec une bonne alliance est significatif
- **4. Chez les patients en traitement de dépendances, il y a un lien entre l'alliance thérapeutique et la pénibilité perçue des contrôles d'urine**
- 5. Les patients en traitement de dépendances qui ont une meilleure alliance trouvent les contrôles d'urine plus pénibles
- 6. Chez les patients en traitement de dépendances, il n'y a pas de lien entre l'alliance thérapeutique et la pénibilité perçue des contrôles d'urine
- 7. Chez les patients en traitement de dépendances, l'alliance est meilleure concernant les buts que concernant les tâches

Tableau 4  
Corrélations

Variables	Correlation Pearson	Significativité (valeur-p)	Correlation de Spearman-Brown (r <sub>s</sub> )
WAI taux total/VAS	-0.67	< 0.001	-0.69
WAI but/VAS	-0.75	< 0.001	-0.80
WAI tâche/VAS	-0.63	< 0.001	-0.64
WAI lien/VAS	-0.48	0.008	-0.51

# Exercice 3.1.2 - Lire et interpréter une corrélation et un OR présentés dans un article sous forme de tableau



## Compliance Differences Between Patients With Breast Cancer in University and County Hospitals

Ian K. Komenaka,<sup>1</sup> Robert E. Pennington, Jr.,<sup>1</sup> Bryan P. Schneider,<sup>2</sup> Chiu-Hsieh Hsu,<sup>3</sup> Laura E. Norton,<sup>1</sup> Susan E. Clare,<sup>1</sup> Noelia M. Zork,<sup>1</sup> Robert J. Goulet, Jr.<sup>1</sup>

Variable	Compliers, N (%)	Crude OR (95% CI)	P Value	Adjusted OR <sup>a</sup> (95% CI)	P Value
<b>Hospital</b>					
County	65 (69.9)	1.00	.0005	1.00	.0020
University	121 (89)	3.47 (1.73-6.97)		10.10 (2.32-43.90)	
<b>Insurance</b>					
Underinsured	71 (75.5)	1.00	.0850	1.00	.7291
Commercial	115 (85.2)	1.86 (0.95-3.63)		1.26 (0.34-4.75)	

<sup>a</sup>Adjusted for hospital, age, insurance, size, stage, ER status, and grade.

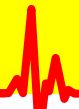
*Predictors of compliance with recommended chemotherapy = prédicteurs de l'adhérence à la chimiothérapie recommandée*

*Compliers = patients adhérents, Crude OR = OR brut; Adjusted OR = OR ajusté*

*Hospital = hôpital, County = régional, University = universitaire,*

*Insurance = assurance, Underinsured = sous-assuré, Commercial = assurance privée*

# Exercice 3.1.3 - Lire et interpréter un t-test présenté dans un article sous forme de tableau



## Amélioration du sentiment de compétence des infirmières pour le dépistage du délirium en médecine générale : une étude quasi-expérimentale

Improvement in perception of nurse's competence in screening for delirium : a quasi-experimental study

### Maryse GRÉGOIRE

Inf. M.A. (généraliste), Conseillère cadre clinicienne  
Centre hospitalier universitaire de Sherbrooke, Canada  
Professeure associée à l'École des sciences infirmières de l'Université de Sherbrooke, Québec, Canada

### Cécile MICHAUD

Inf. Ph.D. (sciences infirmières), Professeure titulaire  
École des sciences infirmières de la Faculté de médecine et des sciences de la santé de l'Université de Sherbrooke, Canada

### Martine PROULX

Inf. Ph.D. (sciences infirmières), Conseillère cadre clinicienne  
Centre hospitalier universitaire de Sherbrooke, Canada

### Patrice LAPLANTE

Ph.D., CCFP (R., Sc.C.)  
Membre du département de médecine générale  
Centre hospitalier universitaire de Sherbrooke, Canada

*GC = Groupe Contrôle, GE = Groupe expérimental (a reçu une formation), ET = écart-type*

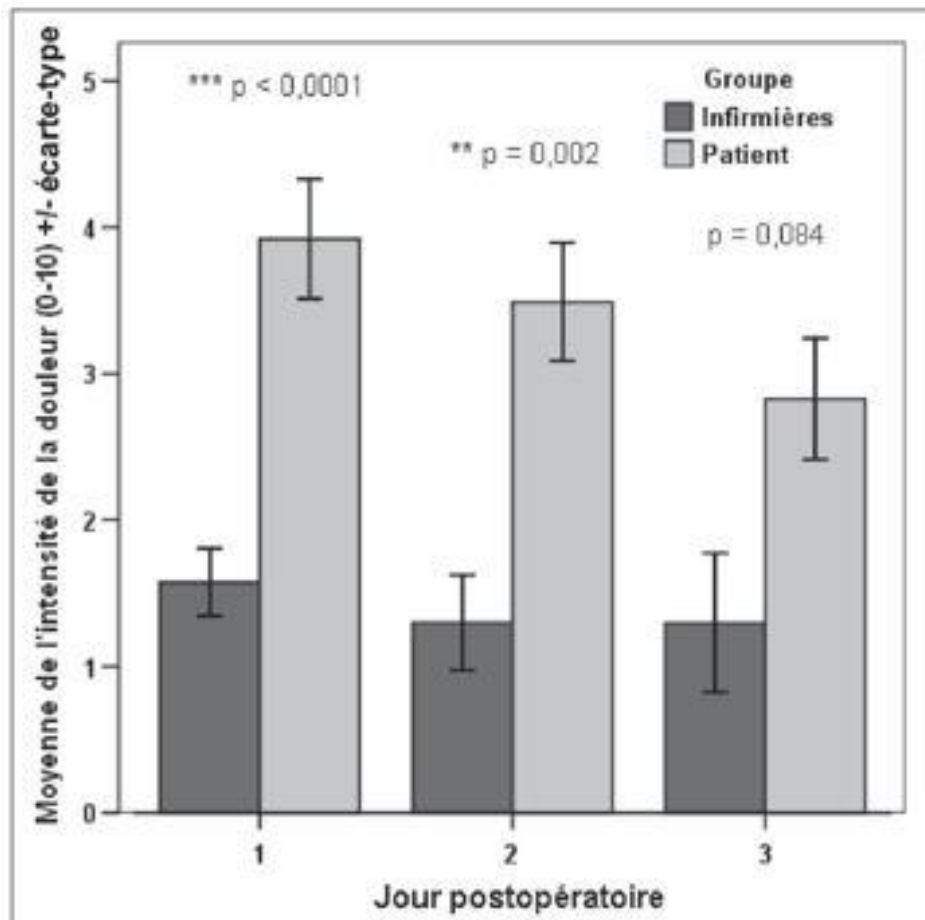
## ■ Sentiment de compétence à dépister le délirium

Avant la formation, les deux groupes ne se distinguaient pas quant à leur sentiment de compétence à dépister le délirium (GE=46,6, GC=47,7). Le groupe expérimental a fait un bond après la formation (60,28; 8,4 ET;  $t=6,29$ ;  $p<0,001$ ) alors que le groupe de comparaison n'a pas bougé (47,0; 8,7 ET;  $t=0,23$ ;  $p=.82$ ). De façon plus générale, le groupe expérimental a présenté des différences statistiquement significatives pour 60% des questions (Encadré 2).

- A quelles hypothèses de recherche ou alternatives pourraient se référer les étoiles indiquant une valeur p qui figurent dans le tableau pour la ligne « Education » ? ATTENTION: nous cherchons toutes les hypothèses possibles qui auraient pu amener à faire les tests statistiques présentés ici, indépendamment qu'elles soient confirmées ou non. Veuillez choisir au moins une réponse :
- 1. Le pourcentage de personnes âgées consommant des médicaments ASH diffèrent entre ceux qui ont un bas et un haut niveau d'éducation
- 2. Les non-consommateurs de médicaments ASH sont plus nombreux que les consommateurs
- 3. Les personnes âgées qui consomment des médicaments ASH ont fait en moyenne moins d'années d'études que celles qui n'en consomment pas
- 4. Le fait de consommer ou non des médicaments ASH prédit le niveau d'éducation des personnes âgées
- 5. Les personnes âgées qui ont fait plus d'années d'études consomment des doses plus élevées de médicaments ASH
- 6. Les personnes âgées qui consomment des médicaments ASH diffèrent de celles qui n'en consomment pas concernant leur nombre moyen d'années d'études



## Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Quel semble décrire ce pattern entre le facteur groupe, et le facteur jour post-op?



Graphique 1) Comparaison de la moyenne de l'intensité de la douleur notée par les patients dans leur journal selon une échelle numérique de 0 à 10 et de la moyenne de l'intensité de la douleur documentée par les infirmières selon une échelle numérique de 0 à 10 sur une période de 24 heures pendant les trois premiers jours postopératoires

Selon les résultats du graphique 1, lorsque les infirmières évaluent et documentent la douleur à l'aide d'une échelle numérique, la moyenne d'intensité de la douleur de leurs évaluations numériques diffère de la moyenne de l'intensité de la douleur des évaluations faites par les patients dans leur journal. Par exemple, pour une même heure, l'infirmière peut noter la douleur à 2 sur 10 alors que le patient l'a évaluée au même moment à 4 sur 10. Pendant la première journée postopératoire, cette différence est très hautement significative ( $p < 0,01$ ), avec une moyenne d'intensité de la douleur documentée par les infirmières de 1,6 ( $\pm 0,2$ ) sur une échelle numérique de 0 à 10, tandis que la moyenne d'intensité documentée par les patients est de 3,8 ( $\pm 0,4$ ). Cet écart tend à diminuer lors des jours postopératoires 2 et 3, mais reste hautement significatif lors de la deuxième journée ( $p < 0,01$ ). L'écart est cependant non significatif ( $p = 0,084$ ) pendant la troisième journée. Il y a donc une différence significative entre la moyenne des évaluations de la douleur documentées par l'infirmière et la moyenne de celles notées par les patients dans le journal pendant chacune des deux premières journées postopératoires.



# Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Commenter les tailles d'effets



Psycho-Oncol. (2011) 5:127-136  
DOI 10.1007/s11839-011-0321-7

ARTICLE ORIGINAL / ORIGINAL ARTICLE DOSSIER

## Le stress au travail chez les infirmières en soins palliatifs de fin de vie selon le milieu de pratique

Stress at work in nurses providing end-of-life palliative care according to care settings

L. Fillion · J.-F. Desbiens · M. Truchon · C. Dallaire · G. Roch

**Tableau 2** Description du stress au travail et analyse de variance pour comparer le stress des infirmières selon le milieu de pratique (CSSS et CH) dans le cadre de l'étude 1 ( $n = 209$ ) et de l'étude 2 ( $n = 751$ )

Variables de stress	Étendue	Total	CH	CSSS	$t^a$	$p$	$d$ de Cohen <sup>b</sup>
		$m$ (ET)	$m$ (ET)	$m$ (ET)			
Étude 1							
Demandes							
Demandes psychologiques	9–36	25,06 (4,20)	26,28 (4,23)	23,95 (3,86)	4,16	0,000	0,55
Efforts	6–24	18,49 (3,50)	18,59 (3,31)	18,39 (3,67)	0,40	0,687	0,06
Ressources							
Latitude décisionnelle	24–96	72,74 (9,98)	71,04 (10,60)	74,30 (9,16)	-2,38+	0,018	0,33
Récompenses	11–44	31,24 (5,02)	31,22 (5,44)	31,26 (4,63)	-0,05	0,958	0,01
Satisfaction au travail	1–7	4,96 (1,09)	4,66 (1,13)	5,23 (0,98)	-3,87*	0,000	0,52
Étude 2							
Demandes							
Demandes psychologiques	9–36	28,21 (3,78)	29,28 (3,68)	26,80 (3,43)	9,36*	0,000	0,66
Efforts	6–24	18,37 (3,11)	18,96 (3,11)	17,57 (2,94)	6,19*	0,000	0,45
Ressources							
Latitude décisionnelle	24–96	73,17 (8,55)	72,60 (8,28)	73,68 (8,85)	-1,70	0,089	0,13
Récompenses	11–44	31,94 (4,79)	32,32 (4,71)	31,43 (4,84)	2,53+	0,012	0,19
Satisfaction au travail	1–7	4,96 (1,10)	4,94 (1,04)	4,99 (1,12)	-0,611	0,541	0,05

CH : centre hospitalier ; CSSS : centres de santé et de services sociaux ; ET : écart-type ; \* :  $p < 0,01$ .

<sup>a</sup> Un seuil significatif de 0,01 a été retenu afin de contrôler l'erreur de type I (approche de Bonferroni).

<sup>b</sup> Taille d'effet.

# Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Commenter les tailles d'effets

Psycho-Oncol. (2011) 5:127-136  
DOI 10.1007/s11839-011-0321-7

ARTICLE ORIGINAL / ORIGINAL ARTICLE

DOSSIER

## Le stress au travail chez les infirmières en soins palliatifs de fin de vie selon le milieu de pratique

Stress at work in nurses providing end-of-life palliative care according to care settings

L. Fillion · J.-F. Desbiens · M. Truchon · C. Dallaire · G. Roch

**Tableau 3** Analyse de variance pour comparer le stress des infirmières selon le milieu de pratique (CSSS et CH — trois secteurs) dans le cadre de l'étude 2 ( $n = 751$ )

Variables de stress	Étendue	CSSS (domicile)		CH Secteurs		$F^a$	$p$	$\omega^b$
		Soins critiques	Soins oncologiques	Soins palliatifs				
Demandes								
Demandes psychologiques	9–36	26,80 (3,47)	30,15 (3,22)	29,11 (3,30)	26,21 (4,09)	54,82*	0,000	0,18
Efforts	6–24	17,55 (2,95)	19,50 (2,95)	18,88 (3,01)	17,24 (3,33)	24,38*	0,000	0,09
Ressources								
Latitude décisionnelle	24–96	73,53 (8,63)	72,20 (8,20)	73,40 (8,89)	74,58 (8,68)	1,92	0,126	0,00
Récompenses	11–44	31,37 (4,81)	31,95 (4,82)	32,85 (4,48)	33,01 (4,70)	4,13*	0,006	0,01
Effets du stress								
Satisfaction au travail	1–7	4,98 (1,13)	4,86 (1,06)	4,99 (0,92)	5,21 (1,00)	1,98	0,115	0,00

CH : centre hospitalier ; CSSS : centres de santé et de services sociaux ; ET : écart-type ; \* :  $p < 0,01$ .

<sup>a</sup> Un seuil significatif de 0,01 a été retenu afin de contrôler l'erreur de type I (approche de Bonferroni).

<sup>b</sup> Coefficient de détermination oméga.

# Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Commenter les tailles d'effets

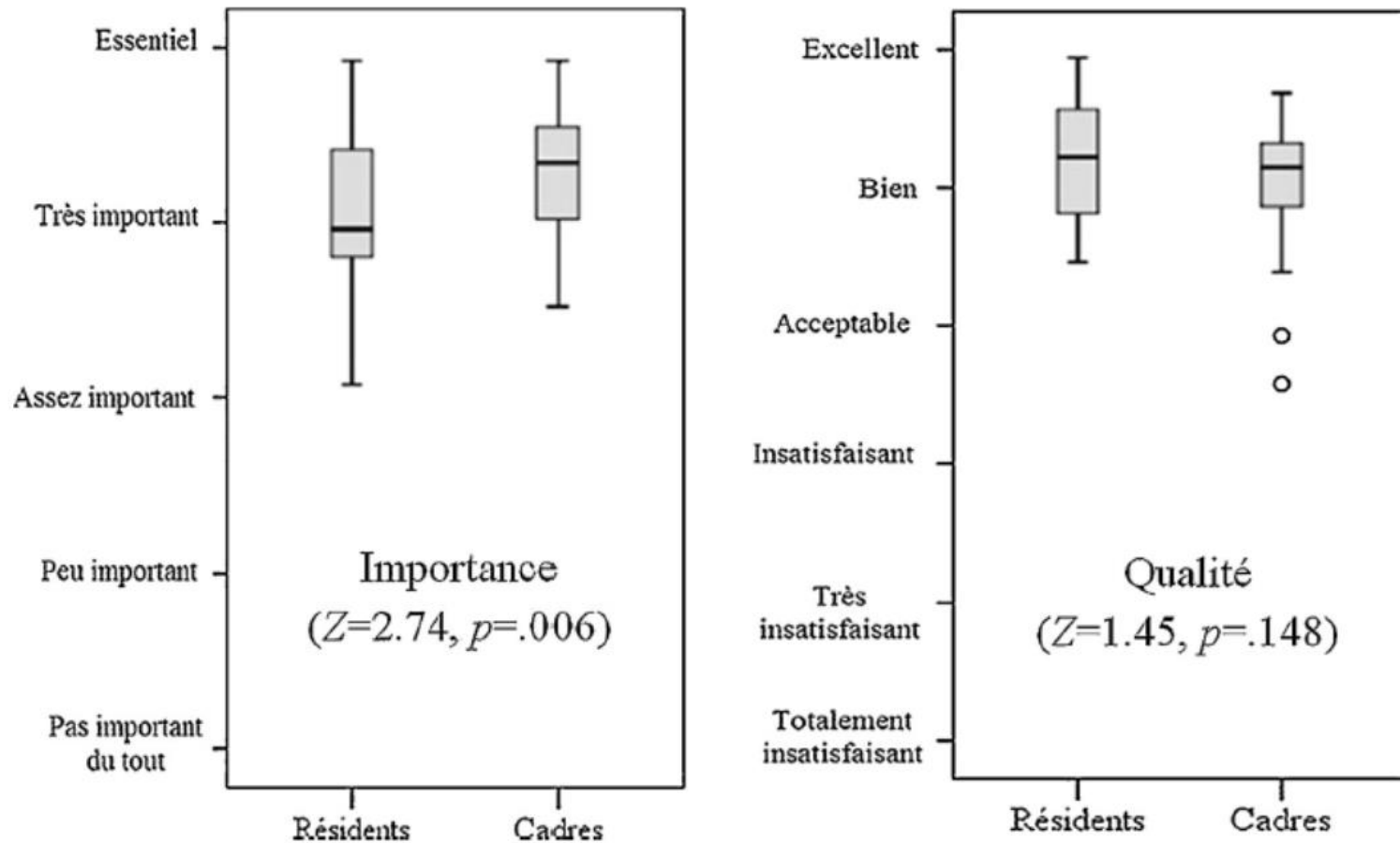


Figure 1. Réponses des cadres et des résident-e-s sur les 27 points évalués (en moyenne, pour les scores bruts), concernant l'importance et la qualité.

# Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Commenter les tailles d'effets



---

## ORIGINAL ARTICLE

---

### Étude descriptive du processus d'évaluation et de documentation de la douleur postopératoire dans un hôpital universitaire

Dave A Bergeron inf BSc<sup>1</sup>, Geneviève Leduc MSc<sup>2</sup>, Serge Marchand PhD<sup>3</sup>, Patricia Bourgault inf PhD<sup>1</sup>

---

DA Bergeron, G Leduc, S Marchand, P Bourgault. Étude descriptive du processus d'évaluation et de documentation de la douleur postopératoire dans un hôpital universitaire. *Pain Res Manage* 2011;16(2):81-86.

Descriptive study of the postoperative pain assessment and documentation process in a university hospital

#### **TABLEAU 4**

**Corrélation entre la moyenne de l'intensité de la douleur notée par les patients dans leur journal selon une échelle numérique de 0 à 10 et celle documentée par les infirmières selon une échelle numérique de 0 à 10 sur une période de 24 heures pendant les trois premiers jours postopératoires**

<b>Journée postopératoire</b>	<b>Nombre de comparaison</b>	<b>r</b>	<b>p</b>
Jour 1	32	0,212	0,243
Jour 2	27	-0,222	0,266
Jour 3	16	0,371	0,157



# Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Quelle stat?

## Adaptation des aînés à la résidence pour personnes âgées: Activité de loisirs et autodétermination

Emin Altintas, Gaëlle Majchrzak, Claire Leconte, and Alain Guerrien  
*Université Lille Nord de France*

Canadian Journal on Aging / La Revue canadienne du vieillissement 29 (4) : 557-565 (2010)  
doi:10.1017/S0714980810000589

### Résultats

#### *Analyses préliminaires*

Les participants hommes et femmes de l'étude ne différenent aucunement pour le nombre de participations aux activités ( $t_{75} = -0.99$  ; *ns*), l'adaptation à la résidence ( $t_{75} = -0.14$  ; *ns*), la satisfaction de vie ( $t_{75} = 0.62$  ; *ns*), l'indice d'autodétermination vis-à-vis des activités de loisirs ( $t_{75} = -1.91$  ; *ns*) ou les différentes dimensions motivationnelles. Toutefois, une différence est relevée entre hommes et femmes : les femmes sont en moyenne plus âgées ( $t_{75} = -2.90$  ;  $p < 0,01$ ).

# Exercice 3.2 (avec feedback) : Lire et interpréter des résultats d'articles présentés sous forme de tableau, graphique ou texte – Quelle stat?

**Table 2.** *Predictors for Recurrence*

	Recurrent (1)–Fast Recovery (0)			
	N = 244			
	Nagelkerke R <sup>2</sup> = 0.25			
	AUC = 0.75 (CI 95%: 0.69-0.81)			
	B (SE)	P	OR	95% CI
Age	-0.02 (0.01)	.153	0.98	0.95-1.01
Sex	0.39 (0.30)	.198	1.47	0.82-2.64
Male (0) Female (1)				
Duration Acute (0)	1.20 (0.35)	<b>.001</b>	3.31	1.67-6.54
Subacute (1)				
Duration Acute (0) Chronic (1)	2.31 (0.45)	<b>&lt;.001</b>	10.10	4.19-24.34
Work status Off (0)	-1.22 (0.71)	.087	0.30	0.07-1.19
Working (1)				
Work status Off (0)	0.77 (0.92)	.934	1.08	0.18-6.55
Retired (1)				
NRS at baseline	0.07 (0.08)	.331	1.08	0.93-1.25
ODI at baseline	0.05 (0.02)	.015	1.06	1.01-1.10