

LE PROTOCOLE DE L'ÉTUDE



Les Etapes d'une Recherche

1. Préparation de l' étude
2. Le protocole de recherche **écrit**
3. Collection et analyse des données
4. L' analyse des résultats préliminaires
5. Analyse et interprétation des résultats
6. Rédaction de l'étude

Préparation de l'étude

Le but de la préparation:

- formuler une hypothèse scientifique pertinente.
- connaître le domaine
 - résultats attendu
 - notions controversées

En utilisant la documentation bibliographique

Préparation de l'étude

- On formule l' hypothèse de l'étude au début de la recherche, basée sur des observations personnelles sur la documentation bibliographique précédente.
- On formule
 - le but
 - les objectives
- On choisit le type d'étude.

PRÉPARATION DE L'ÉTUDE

Préparation de l'étude

- Le but d'une étude

Definition: une aspect plus générale, à atteindre

- Exemples des formulations:

- Evaluer l'association
 - entre un facteur et une maladie/complication,
 - entre un traitement et le résultat,
 - entre deux facteurs / caractéristiques cliniques/ précliniques
- comparer deux techniques diagnostiques
- comparer deux traitements
 - décrire une maladie
- Ex: Evaluer la relation entre la consommation de l'alcool et le cancer de la langue
- Ex: Evaluer l'efficacité et la sécurité de l'aspirine pour réduire l'apparition des événements cardiovasculaires majeurs pour des sujets diabétiques.
- Ex: Evaluer les qualités diagnostiques des radiographies bite-wing pour identifier le granulome dentaire
- Ex. Décrire la prévalence du carcinome basocelulaire du lèvres

PRÉPARATION DE L'ÉTUDE

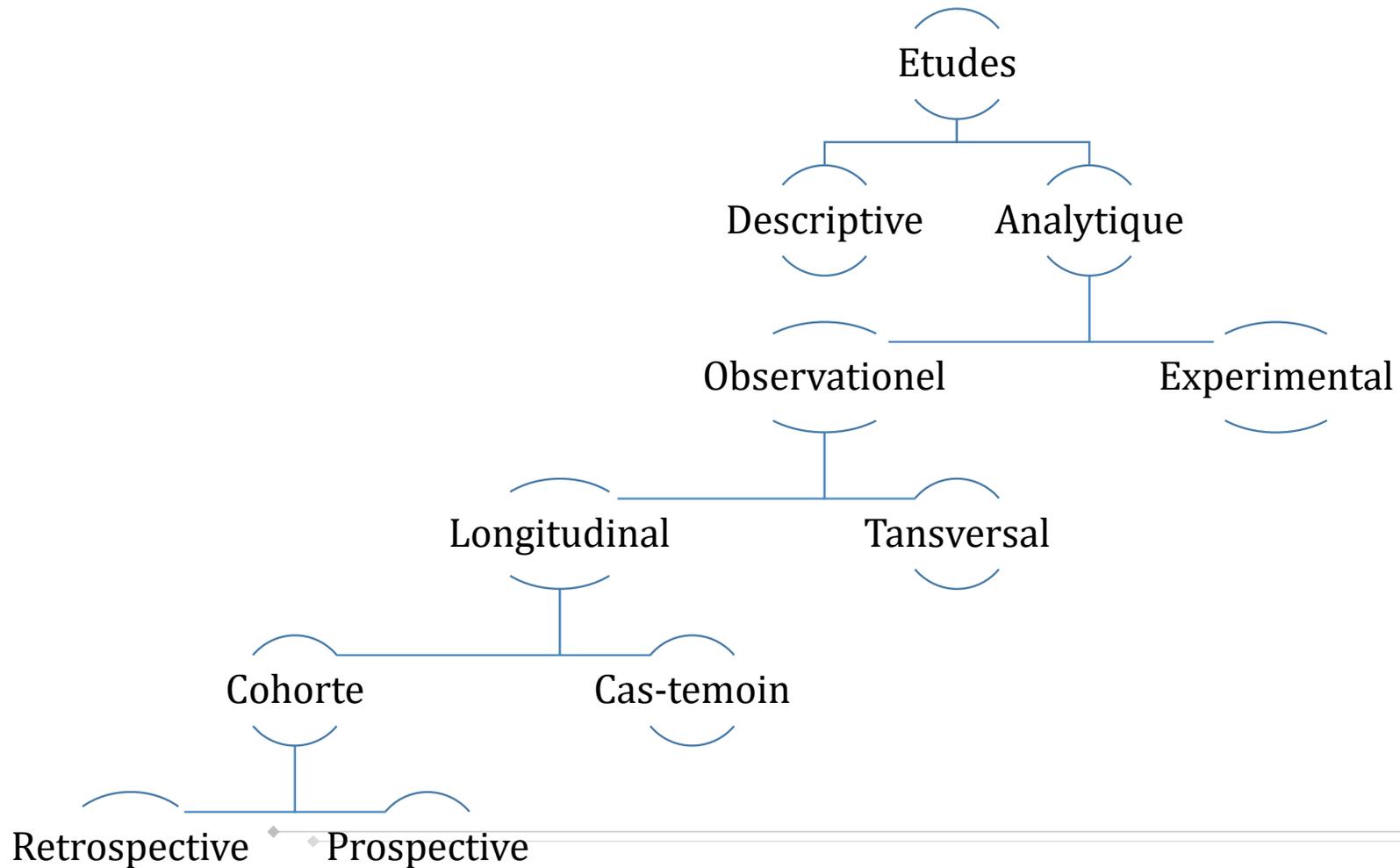
Préparation de l'étude

- Les objectifs d'une étude
- Définition: des aspects plus précis, pratiques, pragmatiques à atteindre
- Types, et exemples:
 - Objectif **majeur**:
 - Tester la différence entre des traitements/ tests diagnostiques/ quantifier la différence
 - Ex. Evaluer l'efficacité de l'aspirine pour réduire l'apparition des accidents cardiovasculaires majeurs pour des sujets diabétiques.
 - Objectifs **secondaires** de l'étude.
 - Autres phénomènes biologiques étudiés dans la même étude,
 - sans altérer la réalisation de l'objectif majeur
 - Autres critères de jugement, critères de jugement plus spécifiques, réactions secondaires
 - Ex. Evaluer la sécurité de l'aspirine en ce qui concerne l'apparition des ulcérations gastriques pour des sujets diabétiques.
 - Ex. Evaluer l'acceptabilité des radiographies bite-wing pour identifier le granulome dentaire

Protocole de l'étude

- On choisit le **domaine de la recherche clinique**
 - Description d'un phénomène de la santé
 - Evaluation des facteurs de risque ou pronostiques
 - Evaluation d'un test diagnostique
 - Evaluation d'un abord thérapeutique
- On choisit le **type d'enquête** (voir diapo suivantes)
 - Descriptive ou analytique
 - Observationnelle ou expérimentale
 - Transversale/longitudinale

Classification selon types d'études



Classification selon les objectifs

- **Enquêtes/ études descriptives**
 - **Definition: une etude sans** des essais ou de **comparaisons** sont effectuées, aucun **lien**, relation, ou association, corrélation, dépendance sont poursuivis
 - **Caractéristiques**
 - destiné à **décrire des phénomènes liés à l'état de santé** d'une population
 - on calcule
 - le pourcentage (d'une maladie (ex. diabète, parodontose) /facteur de risque (tabagisme)),
 - moyenne, déviation standard (ex. de l'âge), médiane, quartiles (ex. du cholestérol, de la profondeur des poches parodontale de la gencive)
 - avec l'intervalle de confiance
 - sans tests statistiques ou autres méthodes statistiques similaires

Classification selon les objectifs

- **Enquêtes/ études descriptives**

- permet **d'évaluer**

- **des index** liés à la **maladie** (ex. infarctus myocardique, cancer de la lèvre)/ **mortalité**
 - de morbidité (concernant les maladies):
 - **l'incidence (nouveau cas dans la population),**
 - **la prévalence (toutes les cas dans une population)**
 - de mortalité (taux des décès dans une population)
 - de létalité (le taux de décès due à une maladie)
- **des statistiques d'exposition à des facteurs** (ex. consommation d'alcool)/ description des certaines **caractéristiques** (ex. âge, sexe, milieu de provenance, cholestérol, hauteur, poids, présence du tartre, fréquence de brossage aux dents, utilisation des solutions de bains de bouche)
 - observe des facteurs
 - à un moment donné,
 - dans l'évolution,
 - leur répartition géographique

Classification selon les objectifs

- **Enquêtes/ études analytiques**

Définition: des études dans lesquels on fait des essais ou de **comparaisons** sont effectuées, des **liens**, relations, ou des associations, ou corrélations ou dépendances sont poursuivis

Types: enquêtes étiologiques et enquêtes d'évaluation (de l'efficacité d'interventions)

Caractéristiques

- on compare des groupes pour **verifier des hypotheses**
- on utilise des tests statistiques / méthodes statistiques inférentiels

Exemples

- Evaluer un lien potentiel entre le facteur pronostique (mauvaise hygiène dentaire) et la maladie (infarctus)
- Tester la différence entre des traitements (comparer aspirine vs. Placebo, pour diminuer l'infarctus; comparer deux matériaux pour obturer des caries dentaires)
- tests diagnostiques (comparer la radiographie avec la tomodensitométrie)

Classification selon la procédure employée

- **Enquêtes/ études d'observation ou observationnelle** (en anglais: observational)
 - **Définition:** des études où les chercheurs n'intervient pas, soit sur les sujets soit sur le cours de la maladie
 - **Types:**
 - descriptives ou analytiques
 - transversales (cross sectional – en anglais)
 - (ex on évalue la relation entre obésité et hypertension a un moment donnée; on évalue la relation entre le tartre dentaire et la sensibilité dentaire au froid)
 - longitudinales
 - cas-témoins (en anglais: case-control)
 - (ex. on compare sujets avec et sans infarctus pour voir la relation avec la situation de la hygiène dentaire)
 - cohorte (en anglais: cohort)
 - (ex. on évalue la consommation des différents aliments ou de l'effort physique sur l'apparition de l'infarctus; on évalue la consommation des différents aliments sur l'apparition du tartre)

Classification selon la procédure employée

- **Enquêtes/ études expérimentales**

Définition: des études où les chercheurs **interviennent** sur les sujets et sur l'évolution de la maladie - par exemple en administrant des traitements (ex aspirine vs. Placebo), interventions chirurgicales (ex appendicectomie), des implants différents (du genou, dentaires), (ou des matériaux d'obturation des caries différents) etc.

Caractéristiques:

- l'observateur **modèle le contexte** de son action - les conditions de l'expérimentation sont **rigoureusement contrôlées**
- la procédure expérimentale est la **plus appropriée** pour déduire des informations de nature étiologique – **on peut tirer des implications de causalité** entre une intervention et le résultat obtenu
- chez l'homme, les études expérimentales en médecine concernent essentiellement les **essais cliniques** (essai randomisé contrôlé)
- Le fait d'utiliser des tests diagnostiques au but diagnostique, n'implique pas une étude expérimentale! mais on peut utiliser un test diagnostique pour voir si ça améliore la santé – et ça est exper.
- Aussi, le fait d'utiliser des tests statistiques n'implique pas une étude expérimentale! On utilise des tests statistiques aussi dans les études observationnelles

Classification selon le temps

- **Enquêtes/ études transversales** (cross sectional - en anglais)

Définition: Des études où les sujets inclus dans l'enquête sont tous les sujets présents **au moment de l'étude**

Caractéristiques:

- On ne suit pas les sujets dans le temps, on les observe une seule fois
- On a des informations sur un seul moment dans le temps
- Représente une photographie d'un phénomène de santé
- la question posée dans les enquêtes transversales: que se passe-t-il à un moment donné ?
- les sujets ne sont pas sélectionnés:
 - ni sur l'exposition comme dans les études de cohortes,
 - ni sur la maladie comme dans les enquêtes cas-témoins

Exemple:

- On évalue la relation entre obésité et hypertension à un moment donné
- On évalue la relation entre la consommation de sucre et la présence de tartre
- On évalue la fréquence de l'obésité des enfants dans l'école

Classification selon le temps

- **Enquêtes/ études Longitudinales**

Definition: des etudes ou on collecte des informations sur les sujets depuis plusieurs moments dans le temps (au moins 2)

Caractéristiques:

- l'organisation est plus difficile
- l'accès a l'information est plus long
- le cout peut être important

Exemple:

- évaluer la relation entre la consommation exagérée du café pendant des décades et l'apparition de l'ostéoporose, ou le changement du couleur des dents
- Evaluer la relation entre l'effort physique et le poids

Avantages	Désavantages
<ul style="list-style-type: none">• des meilleures informations par rapport a des études transversales	<ul style="list-style-type: none">• plus difficile a organiser d'habitude• plus chère d'habitude

Classification selon le temps

Enquêtes/ études **Prospectives** (vers le future)

- Définition: Des études ou l'observation/mesure des caractéristiques se fait après le début de l'étude (on n'a pas des informations depuis le passé) – le chercheur observe/mesure les caractéristiques d'intérêt, d'une manière directe
- Exemple: évaluer la relation entre l'effort physique et la réduction d'anxiété

Avantages

- des **meilleures informations** par rapport a des études rétrospectives et transversales
- **on peut calculer le taux d'incidence ou d'attaque** (+ RIE, RIN, RR, RA, FREE)

Désavantages

- nécessites des **effectifs importants**
- les **perdus de vue** sont une **risque** potentiel pour ce type d' enquêtes (**risque de biais d'attrition**)
- **durée** est d'habitude **longue**
- la **modification dans le temps**
 - des critères **diagnostiques**
 - de l'influence du **facteur** d'exposition
 - sont des **risques potentiels** pour ce type d' enquêtes
- plus **difficile** a **organiser**,
- plus **chère**

Classification selon le temps

Enquêtes/ études Longitudinales rétrospectives (vers le passé)

- Définition: des études ou les observations/mesures des caractéristiques ont été faites dans le passé (avant le début de l'étude)
- Exemple: informations collectées depuis des
- feuilles d'observation,
- bases des données,
- questionnaires/ questions pour le patient/sujet : ex. consommation du café, tabagisme, aliments consommés

Avantages

- ils sont applicable aux **maladies rares** et au **long incubation**
- plus **simples à organiser** d'habitude
- permettent une **logistique plus légère** d'habitude
- permettent un **moindre coût** d'habitude
- durée plus **courte** d'habitude

Désavantages

- risque de
 - biais d'observation
 - biais de mémoire (recall bias)
- on ne peut pas calculer le **taux d'incidence ou d'attaque** (+ RIE, RIN, RR, RA, FREE)

Classification selon les modalités de choix des sujets

- **Enquêtes/ études exhaustives**

Définition: Etudes dans lesquels on étudie toute la population.

La population-cible est la population globale à laquelle doivent s'appliquer les résultats de l'enquête

Avantages

- la **représentativité** des enquêtes exhaustives est **parfaite**

Désavantages

- les enquêtes exhaustives ont une **durée longue** d'habitude
- **logistique/organisation lourde/difficile** d'habitude
- **coûteuses** d'habitude

Classification selon les modalités de choix des sujets

- **Enquêtes/ études par échantillonnage**

Définition: des études où on étudie un échantillon (plusieurs) (on n'étudie pas toute la population)

les sujets sont extraits de la population-cible par **sondage**

Avantages

- plus **simples** à **organiser** d'habitude
- permettent une **logistique** plus **légère** que les enquêtes exhaustives d'habitude
- permettent un **moindre coût** que les enquêtes exhaustives d'habitude
- durée plus **courte** d'habitude

Désavantages

- le résultat n'est qu'une **estimation** de la vraie valeur de la population globale
- il y a des risques de **biais** (de sélection)
- c'est **possible** qu'on **ne peut pas bien généraliser les résultats**

Variables statistiques dans l'étude

La Caractéristique ou la Variable

Types de Variables:

– qualitatives

- dichotomique



- nominale



- ordinale



– quantitatives

- Discrète: 1, 2, 3 ...

- Continue: 63,5 kg; 82,7 kg

- de survie: **survie** (**définition:** variables qui évaluent le temps passé entre l'entrée d'un patient dans un étude et l'apparition d'un événement)
 - le temps passé entre l'entrée d'un patient dans un étude et le décès / guérison / l'apparition d'un signe (métastase, adénopathie, pleurésie, décès), symptôme (la malaise)
 - ex. le temps entre un obturation et la récurrence d'un carie/ la perte de cet dent. Le temps entre la pose d'un implant dentaire et la perte de cet implant



Comment choisir un test statistique pour l'étude - Questions à répondre:

- Quel est le type des variables?
- Combien des échantillons (groups) on a?
- Les échantillons sont:
 - Dépendantes/appariées?
 - (jumeaux/ données répétées/comparaison du partie gauche et droit d'un sujet/études appariées/observation de la même caractéristique avec deux méthodes différentes, voir diapo suivant)
 - Indépendantes?
- Combien des sujets il y a dans les échantillons?
 - Pour les données qualitative:
 - Tableau de contingence: % des cellules théoriques/attendu < 5?
- Quelle est la nature des données?
 - Pour les données quantitatifs:
 - Distribution normale?
 - Variances égales/ inégales?

Comment choisir un test statistique pour l'étude - Questions à répondre:

- **Selon la constitution des échantillons**
 - Pour **échantillons indépendants** (deux ou plusieurs)
 - les observations sont indépendantes à l'intérieur des groupes et d'un groupe à l'autre
 - Ex. on compare des sujets qui ont reçu un traitement hypocholestérolémiant et des sujets qui ont reçu un placebo
 - Pour **échantillons appariés/ dépendants** (deux ou plusieurs)
 - la situation des **mesures répétées** sur les mêmes sujets.
 - ex. on mesure la température d'un patient avant et après la prise d'un médicament qui baisse la température: ibuprofène
 - on compare des **jumeaux identiques**
 - on compare des **échantillons appariés** (pour chaque sujet dans un groupe on trouve un sujet dans l'autre groupe avec caractéristiques similaires)
 - on compare plusieurs observations faites sur certaines parties du corps du même individu, qui conceptuellement ne sont pas différentes – ex: **partie gauche** avec la **partie droite** (ex. un test dermatologique avec un traitement pour diminuer la perte de cheveux). La comparaison ne doit pas être seulement entre la partie gauche et droite. On peut comparer la partie antérieure et postérieure du scalp.
 - on compare **des méthodes** pour **mesurer** une caractéristique (ou deux techniques diagnostiques [ex tension artérielle]) **sur les mêmes sujets**

Comment choisir un test statistique pour l'étude – les plus fréquentes tests:

- Comparaison de données qualitatives
 - **2/ >2 échantillons indépendants (valeurs >5 dans le tableau théorique / entendu pour >80% des cellules)**
 - Test Chi deux (Chi square)
 - **2/ >2 échantillons indépendants (valeurs <5 ou =0 dans le tableau théorique / entendu pour >20% des cellules)**
 - Test exact Fisher
 - **2 échantillons dépendants/appariées**
 - Test Mc Nemar

Comment choisir un test statistique pour l'étude – les plus fréquentes tests:

- Comparaison de données quantitatives qui **suivent une distribution normale**
 - **2 échantillons indépendants**
 - Test Student pour **échantillons indépendants**
 - Avec variances égales / inégales
 - **2 échantillons dépendants/appariés**
 - Test Student pour **échantillons dépendants**
- Comparaison de données quantitatives qui **ne suivent pas une distribution normale**
 - **2 échantillons indépendants**
 - Test Mann Whitney U – Wilcoxon rank sum test
 - **2 échantillons dépendants/appariés**
 - Test Wilcoxon pour **échantillons dépendants** Wilcoxon signed rank test
- Comparaison des données **de survie**
 - **2/ >2 échantillons indépendants**
 - Test Log-Rank

Les hypothèses statistiques dans l'étude

- La création des hypothèses:
 - Certains **questions médicales** ont deux réponses opposées
 - on force beaucoup des questions dans ce format
 - Les réponses correspondent aux deux **modèles** possibles de la réalité
 - Ces deux modèles sont nommés: **hypothèses**
- L'hypothèse **nulle**: H_0
 - **Il n'y a pas** une **différence statistiquement significative** entre 2/ \geq 2 groupes (ex. traitement [ibuprofène vs. placebo] ou facteur de risque [présent vs. absent]) en ce qui concerne la **moyenne/ médiane/ variance/ fréquence ... d'une caractéristique** (ex. résultat du traitement: fréquence de guérison [oui vs. non] ou **moyenne de la température**)
 - **Il n'y a pas** une **relation/lien/association/dépendance/corrélation statistiquement significative** entre 2 caractéristiques/variables: (ex. Facteur de risque [présent vs. absent] – maladie [oui vs. non] , ou traitement [ibuprofène vs. placebo] – résultat du traitement (ex: fréquence de guérison [oui vs. non] ou **moyenne de la température**))
- L'hypothèse **alternative**: H_1 (négation du H_0)
 - **Il y a** une **différence statistiquement significative** entre 2/ \geq 2 groupes en ce qui concerne la **moyenne/ médiane/ variance/ fréquence ... d'une caractéristique**
 - **Il y a** une **relation/lien/association/dépendance/corrélation statistiquement significative** entre 2 caractéristiques/variables

Les hypothèses statistiques dans l'étude

- **C'est très important de préciser:**
 - Ce que vous cherchez: de faire un **comparaison** ou vous cherchez une **lien / relation / association / corrélation / dépendance**
 - quel statistique vous comparez: **des moyennes, des variances, des fréquences**. Ca doit être précisé dans l'énoncé des hypothèses statistiques, et aussi dans l'interprétation du résultat du teste statistique
 - la caractéristique que vous comparez (ex. cholestérol), ou les caractéristiques entre lesquelles vous essayez de trouver une lien/relation/association/ corrélation / dépendance
 - l'expression «statistiquement significative»
 - les populations/groupes a comparer (si il y a)
 - Ou les noms des variables entre lesquelles vous voulez chercher une **lien / relation / association / corrélation / dépendance**

Les hypothèses statistiques dans l'étude

- **La décision avec p-value**

- **La décision avec p-value.**

- **Si $p\text{-value} < (=0,05) \Rightarrow$ on rejete $H_0 \rightarrow$ nous sommes en faveur de H_1**

- Il y a une différence/ Il y a une relation –statistiquement significative

- **Si $p\text{-value} > (=0,05) \Rightarrow$ on NE rejete pas H_0**

- On ne peut pas dire qu' Il y a une différence/ On ne peut pas dire qu' Il y a une relation –statistiquement significative

La présentation des résultats des études

En utilisant:

- **Des estimateurs ponctuels** d'un indicateur médical/statistique d'une caractéristique
 - **Définition:** l'estimateur ponctuel de l'indicateur médical est notre meilleure estimation dans la population cible de la valeur de l'indicateur médical. Il montre aussi ce **qui se passe** chez les patients **dans l'étude**. Il est égale a la valeur de l'indicateur calculée sur les données dans l'étude.
 - Ex.
 - Pourcentage (des diabétiques), prévalence, différence des pourcentages,
 - Moyenne (du poids), médiane, différence des moyennes,
 - OR, RR, RA, Se, Sp, VPP, NNT, ARR, ...
 - OR – il y a 13 fois plus de chance de cancer du poumon pour les fumeurs par rapport a des non fumeurs
- **Des intervalles de confiance 95%**
 - ils présente **la situation dans la population cible**

Les intervalles de confiance dans l'étude

Interprétation statistique de l'intervalle de confiance

- **Interprétation simplifiée/ définition:** l'intervalle de confiance 95% pour un paramètre (OR, RR, RA, Se, Sp, VPP, NNT,ARR, ...) est l'intervalle qui contient le paramètre respectif de la population cible avec une probabilité de 0,95.

En d'autres termes, nous pouvons être sûrs à 95% que l'intervalle contient le paramètre étudié (par exemple, le OR) de la population cible.

Eg.: La vraie valeur de la prévalence du tabagisme dans la population cible se trouve entre 20% et 30% avec une probabilité de 95%.

- **L'interprétation correcte:** l'intervalle de confiance 95% pour un paramètre (OR, par exemple) est l'intervalle qui comprend 95% des paramètres calculés pour aucun nombre d'échantillons qui ont été aléatoirement extraits de la population cible.

Les intervalles de confiance dans l'étude

Interprétation clinique de l'intervalle de confiance:

Est **subjective**, en fonction de

- l'expertise dans le domaine,
- en fonction des résultats des autres études
- En fonction de se qu'on mesure (douleur, anxiété ou décès)
 - Ex: douleur – 3% des sujets avec un douleur réduite n'est pas assez important que 3% des sujets décédées
 - **cliniquement relation relatif importante**
 - les deux extrémités avec valeur clinique importante
 - ex. pourcentage des décès a 5 années dans un cancer: 50% (95% CI 40%-60%)
- cliniquement relation **relatif moins importante**
 - les deux extrémités avec valeur clinique moins importante
 - ex. pourcentage des décès a 5 années dans un cancer: 0,03% (95% CI 0,02%-0,04%)
- l'importance clinique **n'est pas claire**
 - une extrémité avec une valeur clinique importante d'autre moins importante
 - ex. pourcentage des décès a 5 années dans un cancer: 30,03% (95% CI 0,03%-60,03%)

Les intervalles de confiance dans l'étude

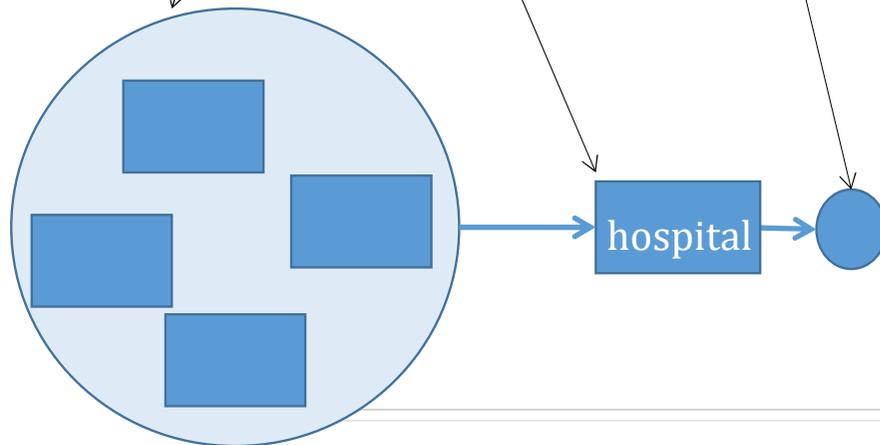
Evaluer la précision d'un résultat – en retardant la largeur de l'intervalle de confiance

- **Plus l'intervalle est large**
la **précision** des résultats est **réduite**
le **nombre des sujets** dans l'étude est **moins grand**
ex. Proportion des fumeurs IC de 5% - 45% (+/- 20% marge d'erreur)
- **Plus l'intervalle est étroit**
la **précision** des résultats **augmente**
le **nombre des sujets** dans l'étude est **plus grand**
ex. Proportion des fumeurs IC de 21% - 28% (+/-3% marge d'erreur)

PLAN DE COLLECTION DES DONNÉES

- Préciser la **population cible** de l'étude =
= la population pour laquelle on pourra appliquer les résultats de l'étude
- Préciser la **population accessible:**

- L'échantillon



PLAN DE COLLECTION DES DONNÉES

- Population cible:
 - **Caractéristiques cliniques** (par exemple, la maladie, le stade de la maladie, les complications, le statut fonctionnel)
 - **Caractéristiques démographiques** (restrictions concernant l'âge, le sexe, le statut socio-économique des patients): nouveau-nés
- Population accessible (en raison de certaines contraintes géographiques ou liées au temps soit pour les sujets ou le chercheur)
 - le hôpital, le cabinet de médecin de famille, une école, ...

PLAN DE COLLECTION DES DONNÉES

- L'échantillon étudié (identifier et noter des critères d'inclusion et d'exclusion possibles, pour vous assurer que il n'a pas de différences notables existent entre les groupes, à l'exception de la différence imposée par votre mode de groupement des sujets)
- Les critères d'inclusion (identifier précisément les caractéristiques des sujets étudiées; si les sujets sont choisis pour former plusieurs groupes, définir des critères distinctifs pour chaque groupe de sujets):
 - **cliniques** (p.ex. que des sujets avec une certaine maladie X en phase Y et des handicaps fonctionnels Z):
 - pour les cas:
 - pour les contrôles:
 - **démographiques** (p.ex. que des sujets âgés de plus de 65 ans, du milieu rural, ayant un faible revenu et l'enseignement primaire au plus):

PLAN DE COLLECTION DES DONNÉES

- L'échantillon étudié:
- **Les critères d'exclusion** (*appliquées uniquement aux sujets qui se conforment aux critères d'inclusion ci-dessus ; si pas relevant, tout ou partie des critères d'exclusion peuvent être omises*):
 - Facteurs de biais (*par exemple des maladies concomitantes /traitements coexistants*):
 - Effets indésirables:
 - Facteurs qui rendent la collecte de données difficile, voire impossible:
 - des dossiers des patients incomplets.
 - Questions éthiques:

Les types de collection des données selon le regroupement

- Un group (échantillon représentative, cohorte, série des cas)
- Exposée - non exposée
- Cas - témoin

Les types de collection des données selon le regroupement

- a. Un group (échantillon représentatif - la recherche est effectuée sur un échantillon en miroir de toutes les caractéristiques de la population cible à partir de laquelle il a été extrait, ou soit une cohorte des sujets (un group des sujets avec une caractéristique commune, ou les mêmes critères d'inclusion et d'exclusion), ou une série consécutive des cas, ou une série des cas de convenance/pratique... (ex. un group des diabétiques)
- b. Exposé - non exposé
 - Deux groupes sont suivis: l'un exposé, l'autre non exposé à un facteur pronostique (ex. expose – utilise le sucre, non expose – n'utilise pas du sucre)
 - On fixe une facteur d'exposition prédéterminé, et on observe un objective dans les deux groups (l'apparition d'une maladie)
- c. Cas – témoins
 - Deux groupes sont suivis: les cas, qui ont la maladie prédéterminé, et les témoins, qui sont indemnes de la maladie prédéterminée (ex. avec/ sans parodontite)
 - On fixe la maladie, et on observe un objective dans les deux groups (la présence des facteurs de risque)

PLAN DE COLLECTION DES DONNÉES

Choisir **l'échantillon** pour l'étude

- Choisir le type d'échantillonnage:
 - Voir diapo. suivant
- Choisir la modalité d'échantillonnage implique:
 - Calculer la taille de l'échantillon
 - Résoudre les problèmes d'abandon
 - perdus du vue,
 - non-répondeurs
 - Définir les critères
 - d'inclusion
 - d'exclusion des sujets

PLAN DE COLLECTION DES DONNÉES

Pour construire un *échantillon **représentative*** il faut sélectionner:

Ces éléments d'une manière aléatoire

Un nombre suffisant d'éléments.

Si l'échantillon est représentative on peut généraliser les résultats depuis l'échantillon vers la population cible, sinon, on peut pas

Types d'échantillonnage:

- *Échantillonnage probabiliste*
 - aléatoire simple
 - stratifié
 - systématique
 - avec clusters
- *Échantillonnage non probabiliste*
 - pratique/convenable
 - consécutive

PLAN DE COLLECTION DES DONNÉES

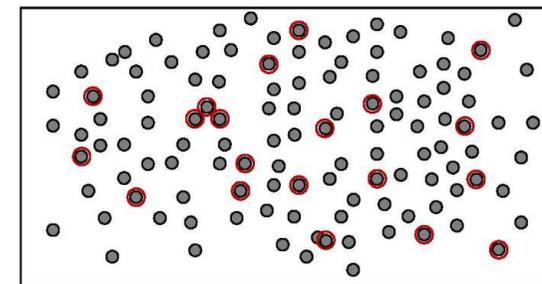
Échantillonnage probabiliste

Échantillonnage aléatoire simple

Définition: échantillon obtenu par l'extraction aléatoire (randomise) d'éléments d'une population.

Méthodes: On peut utiliser

- des listes des nombres aléatoires – tableaux dans des livres de statistique,
- des programmes sur l'ordinateur pour générer la liste
 - (e.g. fonction RAND dans Excel, ou Random Number Generator dans Analysis, dans menu Tools)
- des outils sur internet: randomizer.org)



PLAN DE COLLECTION DES DONNES

Échantillonnage probabiliste

Échantillonnage aléatoire simple

- randomizer.org:

RESULTS

PRINT

DOWNLOAD

CLOSE

1 Set of 2,000 Unique Numbers

Range: From 1 to 2,000— Sorted from Least to Greatest

Set #1

p1=1, p2=2, p3=3, p4=4, p5=5, p6=6, p7=7, p8=8, p9=9, p10=10, p11=11, p12=12, p13=13, p14=14, p15=15, p16=16, p17=17, p18=18, p19=19, p20=20, p21=21, p22=22, p23=23, p24=24, p25=25, p26=26, p27=27, p28=28, p29=29, p30=30, p31=31, p32=32, p33=33, p34=34, p35=35, p36=36, p37=37, p38=38, p39=39, p40=40, p41=41, p42=42, p43=43, p44=44, p45=45, p46=46, p47=47, p48=48, p49=49, p50=50, p51=51, p52=52, p53=53, p54=54, p55=55, p56=56, p57=57, p58=58, p59=59, p60=60, p61=61, p62=62, p63=63, p64=64, p65=65, p66=66, p67=67, p68=68, p69=69, p70=70, p71=71, p72=72, p73=73, p74=74, p75=75, p76=76, p77=77, p78=78, p79=79, p80=80, p81=81, p82=82, p83=83, p84=84, p85=85, p86=86, p87=87, p88=88, p89=89, p90=90, p91=91, p92=92, p93=93, p94=94, p95=95, p96=96, p97=97, p98=98, p99=99, p100=100, p101=101, p102=102, p103=103, p104=104, p105=105, p106=106, p107=107, p108=108, p109=109, p110=110, p111=111,

GENERATE NUMBERS

How many sets of numbers do you want to generate?

1

▶ Help

How many numbers per set?

2000

▶ Help

Number range (e.g., 1-50)

1

2000

▶ Help

Do you wish each number in a set to remain unique?

Yes

▶ Help

Do you wish to sort the numbers that are generated?

Yes, least to greatest

▶ Help

How do you wish to view your random numbers?

Place Markers Within

▶ Help

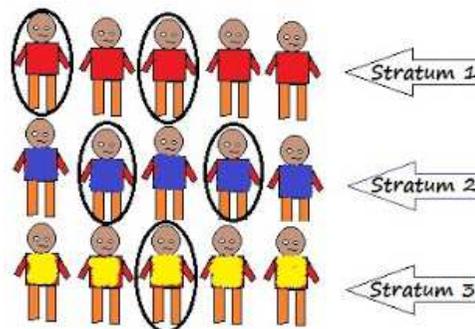
RANDOMIZE NOW!

PLAN DE COLLECTION DES DONNÉES

Échantillonnage probabiliste

Échantillonnage Aléatoire stratifié

Définition: échantillon obtenu par l'extraction aléatoire (randomise) d'éléments d'une population, mais sur des strates avec certain proportion.



Exemple:

Si on sait que le pourcentage des sujets dans le milieu urbain est 70%, et 30% pour milieu rural, on va extraire dans un manière aléatoire, des sujets jusqu'a on va trouver 70% des sujets dans le milieu urbain.

Si on trouve plus des sujets dans le milieu urbain on va ignorer un parti d'eux pour trouver le pourcentage des sujets dans le milieu rural qu'on est intéressé

ÉCHANTILLONNAGE

Échantillonnage probabiliste

Aléatoire stratifié

> *Vaccine*. 2021 Jan 8;39(2):247-254. doi: 10.1016/j.vaccine.2020.12.009. Epub 2020 Dec 5.

Individual preferences for COVID-19 vaccination in China

Anli Leng ¹, Elizabeth Maitland ², Siyuan Wang ³, Stephen Nicholas ⁴, Rugang Liu ⁵, Jian Wang ⁶

Affiliations + expand

PMID: 33328140 PMCID: PMC7719001 DOI: 10.1016/j.vaccine.2020.12.009

[Free PMC article](#)

Abstract

Background: Vaccinations are an effective choice to stop disease outbreaks, including COVID-19. There is little research on individuals' COVID-19 vaccination decision-making.

Objective: We aimed to determine individual preferences for COVID-19 vaccinations in China, and to assess the factors influencing vaccination decision-making to facilitate vaccination coverage.

Methods: A D-efficient discrete choice experiment was conducted across six Chinese provinces selected by the stratified random sampling method. Vaccine choice sets were constructed using seven attributes: vaccine effectiveness, side-effects, accessibility, number of doses, vaccination sites, duration of vaccine protection, and proportion of acquaintances vaccinated. Conditional logit and latent class models were used to identify preferences.

Results: Although all seven attributes were proved to significantly influence respondents' vaccination decision, vaccine effectiveness, side-effects and proportion of acquaintances vaccinated were the most important. We also found a higher probability of vaccinating when the vaccine was more effective; risks of serious side effects were small; vaccinations were free and voluntary; the fewer the number of doses; the longer the protection duration; and the higher the proportion of acquaintances vaccinated. Higher local vaccine coverage created altruistic herd incentives to vaccinate rather than

Using a stratified random sampling method, six Chinese provinces were chosen based on high, medium and low GDP per capita, geographically covering east, central and west regions of China. In each province, three cities were chosen, also divided into high, medium and low GDP per capita, with 314 individuals randomly surveyed in each city.

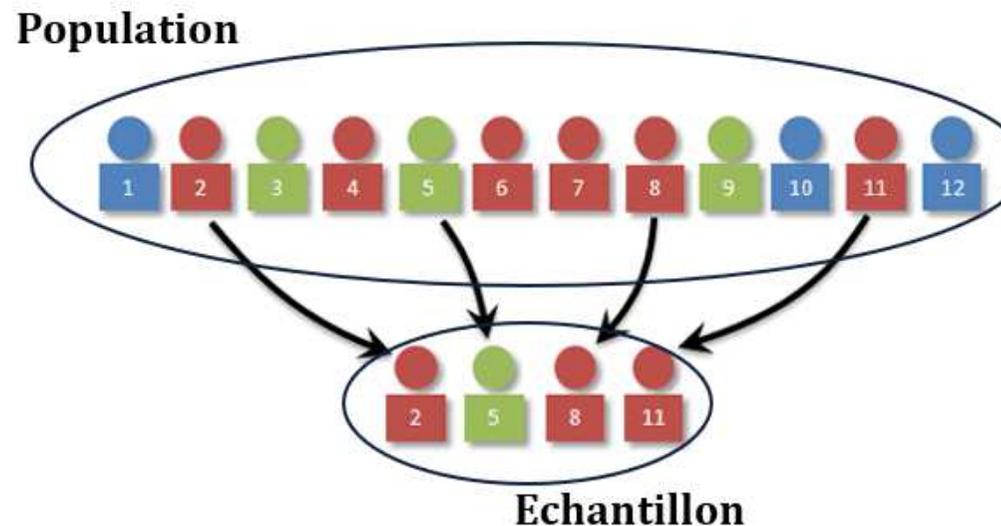
PLAN DE COLLECTION DES DONNÉES

Échantillonnage systématique

Définition: un échantillon obtenu par sélection de chaque k-eme élément

e.g. on utilise la liste des patients inscrits dans le cabinet d'un médecin.

Si on veut un échantillon de 200 sujets, et la liste a 600 sujets, on va extraire chaque 3eme sujet



ÉCHANTILLONNAGE

systematique

> [Pan Afr Med J. 2021 May 13;39:38. doi: 10.11604/pamj.2021.39.38.27648. eCollection 2021.](#)

Availability of PPEs and training status of health professionals on COVID-19 in Silte Zone, Southern Ethiopia

Mohammed Muze ¹, Bahredin Abdella ¹, Abdilmejid Mustefa ¹, Abas Ali ¹, Abdulfeta Abdo ², Abas Lukman ², Abdulfeta Shafi ³, Shukure Uomer ³, Yesufe Badege ⁴, Abdilmejid Mutteba ⁴, Bayesa Tolasa ⁵, Sister Hossae ⁵, Sultan Shukur ⁶, Ebrahim Ahmed ⁷, Abdu Kemal ⁸, Tadela Erena ⁹

Affiliations + expand

PMID: 34422161 PMID: PMC8356931 DOI: 10.11604/pamj.2021.39.38.27648

[Free PMC article](#)

Abstract

Introduction: recent infectious disease outbreaks like COVID-19 highlights the importance of personal protective equipments and competent professionals on public health preparedness and response in health care systems. Hence, understanding availability of personal protective equipments and training status of health professionals is very important to fill the gap of COVID-19 preparedness and response. Therefore, this study was conducted to assess availability and adequacy of personal protective equipments and health professional's training status on COVID-19 in Silte Zone, southern Ethiopia.

Methods: cross sectional study was conducted from August to October 2020 in Silte Zone. First four weredas from 13 were selected randomly. **Systematic sampling** technique was used to select 351 health professionals from 13 health facilities of selected weredas.

Results: overall, only 36.1% of the health professionals have received adequate training on COVID-19. About 30% of the health professionals had taken training on emergency plan of COVID-19, about 33% had been taught on COVID-19 treatment procedures. Majority 80.9% of the respondents indicated that personal protective equipments are inadequately available. Face masks, hand sanitizers and eye goggles were most scarce PPEs.

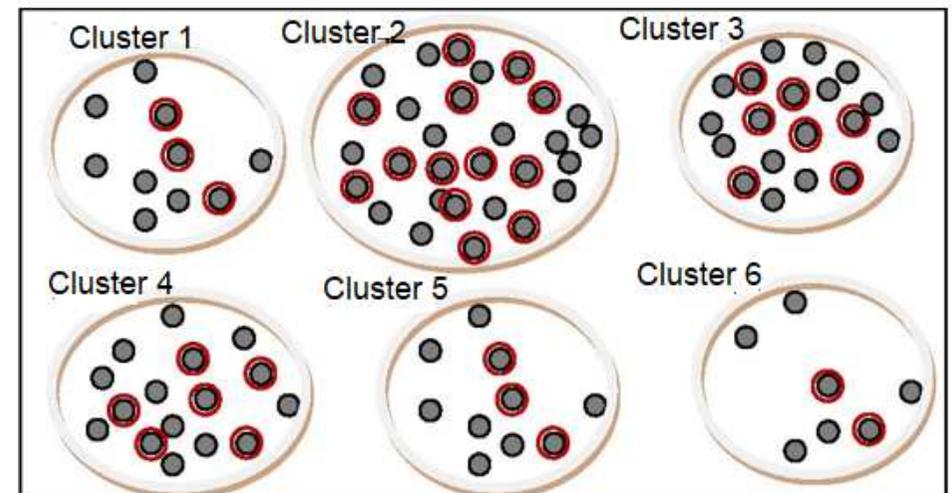
PLAN DE COLLECTION DES DONNÉES

Échantillonnage avec clusters (en grappes)

Définition: échantillon obtenu par l'extraction aléatoire des clusters d'une population.

Un cluster peut être:

- un hôpital,
- un cabinet d'un médecin,
- un école,
- un ville,
- un famille, ...



L'échantillonnage en grappes peut être

Non probabiliste si l'extraction n'est pas aléatoire

Dans chaque cluster on peut après extraire tous les sujets ou on peut utiliser d'autres méthodes d'échantillonnage: aléatoire simple, stratifié, systématique,

Échantillonnage

- Un échantillonnage en grappes (cluster – en anglais) – pour plusieurs hôpitaux c'est plus bon que l' échantillonnage dans un seul grappe/hôpitaux/cabinet...
- en grappes (cluster – en anglais) – pour plusieurs hôpitaux, et a l' intérieur de chaque hôpital on peut utiliser systématique, aléatoire simple, stratifié

ÉCHANTILLONNAGE

En grappes (clusters – en anglais)

> [Psychiatry Res. 2020 May;287:112934. doi: 10.1016/j.psychres.2020.112934. Epub 2020 Mar 20.](#)

The psychological impact of the COVID-19 epidemic on college students in China

Wenjun Cao ¹, Ziwei Fang ², Guoqiang Hou ³, Mei Han ², Xinrong Xu ², Jiaxin Dong ², Jianzhong Zheng ⁴

Affiliations + expand

PMID: 32229390 PMID: PMC7102633 DOI: 10.1016/j.psychres.2020.112934

[Free PMC article](#)

Abstract

A COVID-19 epidemic has been spreading in China and other parts of the world since December 2019. The epidemic has brought not only the risk of death from infection but also unbearable psychological pressure. We sampled college students from Changzhi medical college by using **cluster sampling**. They responded to a questionnaire packet that included the 7-item Generalized Anxiety Disorder Scale (GAD-7) and those inquiring the participants' basic information. We received 7,143 responses. Results indicated that 0.9% of the respondents were experiencing severe anxiety, 2.7% moderate anxiety, and 21.3% mild anxiety. Moreover, living in urban areas (OR = 0.810, 95% CI = 0.709 - 0.925), family income stability (OR = 0.726, 95% CI = 0.645 - 0.817) and living with parents (OR = 0.752, 95% CI = 0.596 - 0.950) were protective factors against anxiety. Moreover, having relatives or acquaintances infected with COVID-19 was a risk factor for increasing the anxiety of college students (OR = 3.007, 95% CI = 2.377 - 3.804). Results of correlation analysis indicated that economic effects, and effects on daily life, as well as delays in academic activities, were positively associated with anxiety symptoms ($P < .001$). However, social support was negatively correlated with the level of anxiety ($P < .001$). It is suggested that the mental health of college students should be monitored during epidemics.

ÉCHANTILLONNAGE

Échantillonnage non probabiliste

Le plus utilisé dans les études cliniques

- **pratique/convenable**

- en utilisant les plus convenables sources des sujets, en fonction de leur relative facilité d'accès: volontaires parmi les médecins/étudiants; ou parmi des amis, famille, collègues d'un médecin, ou des patients.
- ca pose des problèmes – les volontaires, et les proches sont parfois différent du reste de la population,
- une étude dans un hôpital sans préciser que c'est une collecte consécutive des sujets, est plus probable un échantillonnage pratique

- **consécutive**

- le **meilleur échantillonnage non probabiliste**
- on prend toutes les patients qui viennent avec les critères d'intérêt pour l'étude

Échantillonnage non probabiliste

> *Neurol Sci.* 2020 Nov;41(11):3057-3061. doi: 10.1007/s10072-020-04731-9. Epub 2020 Sep 19.

Seizure in patients with COVID-19

Amir Emami ¹, Nima Fadakar ², Ali Akbari ³, Mehrzad Lotfi ⁴, Mohsen Farazdaghi ², Fatemeh Javanmardi ¹, Tahereh Rezaei ⁵, Ali A Asadi-Pooya ^{6 7}

Affiliations + expand

PMID: 32949289 PMID: PMC7501768 DOI: 10.1007/s10072-020-04731-9

[Free PMC article](#)

Abstract

Objective: The purpose of the current study was to collect the data on the occurrence of seizures in patients with COVID-19 and to clarify the circumstances of the occurrence of seizures in these patients.

Methods: All consecutive patients who referred to healthcare facilities anywhere in Fars province (located in South Iran with a population of 4.851 million people) from February 19 until June 2, 2020, and had confirmed COVID-19 by positive result on polymerase chain reaction testing and seizure were included.

Results: During the study period, 6,147 people had confirmed COVID-19 in Fars province, Iran; 110 people died from the illness (case fatality rate 1.79%). During this time period, five people had seizures (seizure rate 0.08%). In four patients, seizure was one of the presenting manifestations, and in one person, it happened during the course of hospital admission. Two patients had status epilepticus. All patients experienced hypoxemia and four of them needed respirator. Two patients had related metabolic derangements and one had cerebrospinal fluid (CSF) lymphocytic pleocytosis. Brain imaging was abnormal in three patients. Four patients died.

Conclusion: New-onset seizures in critically ill patients with COVID-19 should be considered as acute symptomatic seizures and the treating physician should try to determine the etiology of the seizure and manage the cause immediately and appropriately. Detailed clinical, neurological, imaging, and electrophysiological investigations and attempts to isolate SARS-CoV-2 from CSF may clarify the role played by this virus in causing seizures in these patients.

> *BMC Public Health.* 2020 Sep 4;20(1):1351. doi: 10.1186/s12889-020-09429-3.

Early effects of the COVID-19 pandemic on physical activity and sedentary behavior in children living in the U.S

Genevieve F Dunton ^{1 2}, Bridgette Do ³, Shirlene D Wang ³

Affiliations + expand

PMID: 32887592 PMID: PMC7472405 DOI: 10.1186/s12889-020-09429-3

[Free PMC article](#)

Abstract

Background: COVID-19 restrictions such as the closure of schools and parks, and the cancellation of youth sports and activity classes around the United States may prevent children from achieving recommended levels of physical activity (PA). This study examined the effects of the COVID-19 pandemic on PA and sedentary behavior (SB) in U.S. children.

Method: Parents and legal guardians of U.S. children (ages 5-13) were recruited through convenience sampling and completed an online survey between April 25-May 16, 2020. Measures included an assessment of their child's previous day PA and SB by indicating time spent in 11 common types of PA and 12 common types of SB for children. Parents also reported perceived changes in levels of PA and SB between the pre-COVID-19 (February 2020) and early-COVID-19 (April-May 2020) periods. Additionally, parents reported locations (e.g., home/garage, parks/trails, gyms/fitness centers) where their children had performed PA and their children's use of remote/streaming services for PA.

Results: From parent reports, children (N = 211) (53% female, 13% Hispanic, M_{age} = 8.73 [SD = 2.58] years) represented 35 states and the District of Columbia. The most common physical activities during the early-COVID-19 period were free play/unstructured activity (e.g., running around, tag) (90% of children) and going for a walk (55% of children). Children engaged in about 90 min of school-related sitting and over 8 h of leisure-related sitting a day. Parents of older children (ages 9-13) vs. younger children (ages 5-8) perceived greater decreases in PA and greater increases in SB from the pre- to early-COVID-19 periods. Children were more likely to perform PA at home indoors or on neighborhood streets during the early- vs. pre-COVID-19 periods. About a third of children used remote/streaming services for activity classes and lessons during the early-COVID-19 period.

Conclusion: Short-term changes in PA and SB in reaction to COVID-19 may become permanently entrenched, leading to increased risk of obesity, diabetes, and cardiovascular disease in children. Programmatic and policy strategies should be geared towards promoting PA and reducing SB over the next 12 months.

ÉCHANTILLONNAGE

Échantillonnage non probabiliste

Quand vous ne pouvez identifier quel type d'échantillonnage a été réalisée dans une étude (section Matériel et méthodes) clinique (dans des hôpitaux), il y a plusieurs possibilités:

- Soit l'étude a utilisé une échantillonnage pratique (convenable), et les auteurs ont caché cette information (ou parfois ils ont oublié), ou ils ne savent pas que cette information doit être écrite – surtout dans des articles plus anciennes) depuis l'article, parce que ils savent que méthodologiquement n'est pas la meilleure technique (variante plus probable)
- Soit l'étude a utilisé une échantillonnage consécutive (idéal dans des études cliniques), mais les auteurs ont oublié de le préciser. (variante moins probable)

Exemple: ici on peut pas identifier la technique d'échantillonnage => on pense que il y a une échantillonnage plutôt pratique

Materials and Methods

Subjects

The study was conducted on 150 T2DM patients diagnosed as per American Diabetes Association criteria (age 44.12 ± 8.5 year, M:F ratio 57.33 vs. 42.67). The patients with secondary diabetes mellitus (DM), other endocrinopathy, on drugs such as steroids, Pioglitazone, Metformin and with conditions that could influence insulin resistance, IHD, infections were excluded from the study. Also, T2DM patients who were directly exposed to chemicals or having cancer, renal failure or HIV were excluded. 150 healthy normal glucose tolerant controls (age 44.1 ± 10.4 years, M:F ratio 55.33 vs. 44.67) were studied as controls. The institutional ethics committee of

PLAN DE COLLECTION DES DONNÉES

Le **calcul de taille de l' échantillon étudiée dépend en général** de:

- ce qu'on veut évaluer
- dimension de la différence a observer / précision des résultats
- nature des données
- variabilité des données
- aspects statistiques:
 - niveau de confiance (ex. 95%)
 - niveau de signification statistique alpha (ex. 5%)
 - puissance (80%)
 - utilisation du p bilatéral / unilatéral
 - groupes de dimension égaux / inégaux

PLAN DE COLLECTION DES DONNÉES

Le **calcul de taille de l'échantillon étudiée pour un proportion d'une caractéristique qualitative dichotomique dépend de:**

- précision des résultats (ex. niveau d'erreur de 3%)
- nature des données
 - la valeur de la proportion dans des études antérieurs
 - ou le pire scénario possible – proportion inconnue, on le considère 50%
- aspects statistiques:
 - niveau de confiance (ex. 95%)

<http://www.raosoft.com/samplesize.html>

LE CALCUL DE TAILLE DE L'ÉCHANTILLON ÉTUDIÉE POUR UN PROPORTION

Erreur=5%

What margin of error can you accept? 5% is a common choice	<input type="text" value="5"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="50"/> %
Your recommended sample size is	377

Proportion attendue = 40%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="40"/> %
Your recommended sample size is	975

Erreur=3%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="50"/> %
Your recommended sample size is	1014

Proportion attendue = 30%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="30"/> %
Your recommended sample size is	858

<http://www.raosoft.com/samplesize.html>

LE CALCUL DE TAILLE DE L'ÉCHANTILLON ÉTUDIÉE POUR UN PROPORTION

Proportion attendue = 10%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="10"/> %
Your recommended sample size is	377

Proportion attendue = 90%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="90"/> %
Your recommended sample size is	377

Proportion attendue = 25%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="25"/> %
Your recommended sample size is	770

Proportion attendue = 75%

What margin of error can you accept? 5% is a common choice	<input type="text" value="3"/> %
What confidence level do you need? Typical choices are 90%, 95%, or 99%	<input type="text" value="95"/> %
What is the population size? If you don't know, use 20000	<input type="text" value="20000"/>
What is the response distribution? Leave this as 50%	<input type="text" value="75"/> %
Your recommended sample size is	770

<http://www.raosoft.com/samplesize.html>

PLAN DE COLLECTION DES DONNÉES

Le calcul de taille de l'échantillon étudiée pour un proportion

Si on tiennent toutes les contraintes constantes (eg. alpha), et on augmente l'erreur (on diminue la précision), le nombre des sujets va diminuer

Si tiennent toutes les contraintes constantes (eg. alpha), et on augmente la précision (on diminue l'erreur), le nombre des sujets va augmenter

Si la proportion attendue est 50% le nombre des sujets est le plus grand possible

Plus la proportion attendue e proche a 50%, plus le nombre des sujets est grand

Plus la proportion attendue e proche a 0%, plus le nombre des sujets est petit

Plus la proportion attendue e proche a 100%, plus le nombre des sujets est petit

Des proportion attendue de distance égale a 50% (ex. 10% avec 90% ou 25% avec 75%) ont des nombres des sujets égales

PLAN DE COLLECTION DES DONNÉES

Le **calcul de taille de l'échantillon étudiée** pour **comparer la proportion d'une caractéristique qualitative dichotomique entre deux groupes** dépend de:

- dimension de la différence à observer (ex. 10%)
- nature des données
 - la valeur de la proportion dans le group de control connue depuis des études antérieurs
- aspects statistiques:
 - niveau de signification statistique alpha (ex. 5%)
 - puissance (80%) ($1 - \beta$)
 - utilisation du p bilatéral / unilatéral
 - groupes de dimension égaux / inégaux
 - <http://statpages.org/proppowr.html>

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions: <http://statpages.org/proppowr.html>

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.40	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.60	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

Puissance=80%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	97	97	194
With Continuity Correction:	107	107	213

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	85	(Usually 80)
Group 1 Population Proportion:	.40	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.60	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

Puissance=85%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	111	111	221
With Continuity Correction:	120	120	241

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	90	(Usually 80)
Group 1 Population Proportion:	.40	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.60	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

Puissance=90%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	129	129	259
With Continuity Correction:	139	139	278

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	95	(Usually 80)
Group 1 Population Proportion:	.40	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.60	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

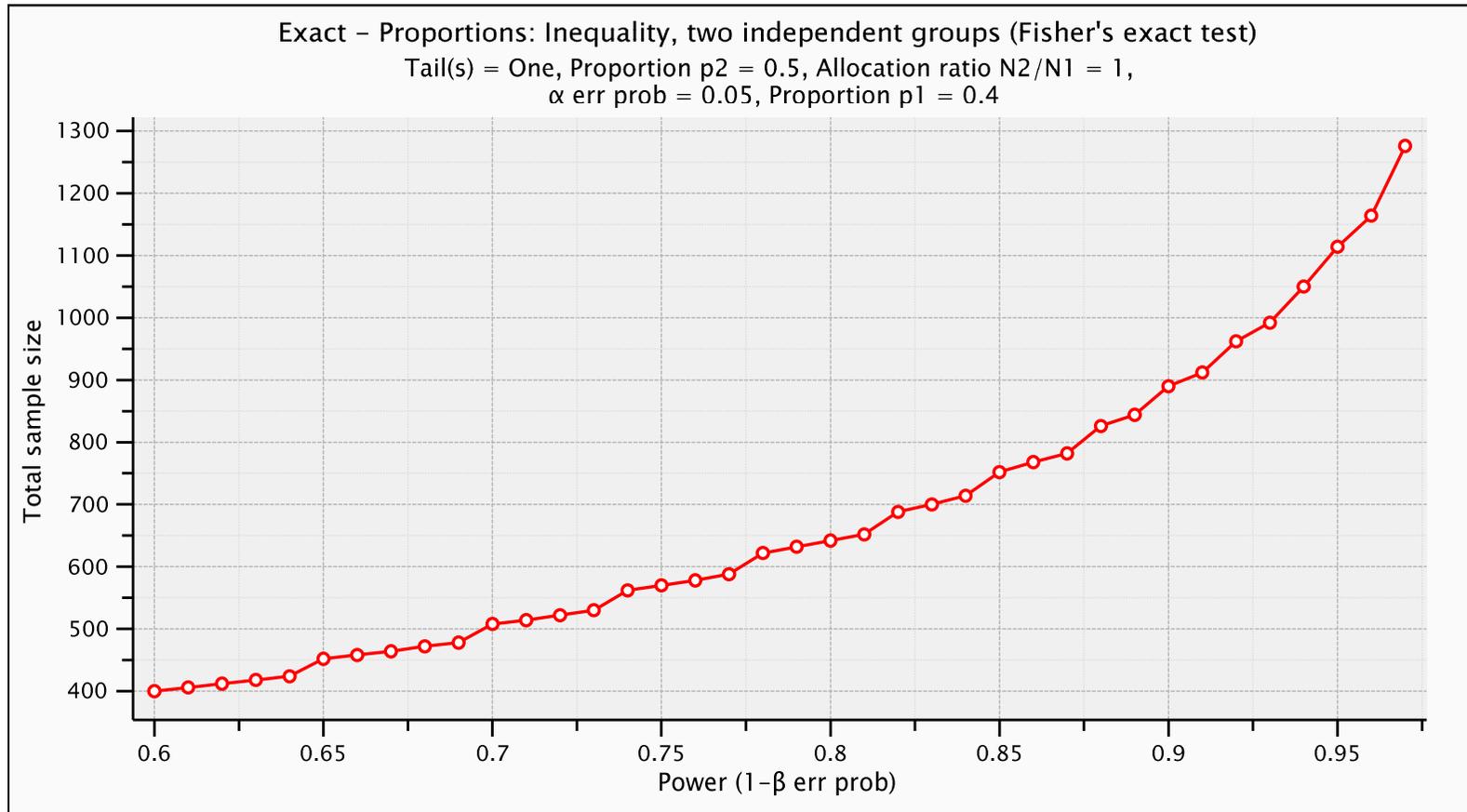
Puissance=95%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	159	159	319
With Continuity Correction:	169	169	339

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions <http://statpages.org/proppowr.html>



Taille totale des 2 échantillons en fonction de la puissance du test

PLAN DE COLLECTION DES DONNÉES

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.30	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.50	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

différence=20%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	93	93	186
With Continuity Correction:	103	103	206

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.40	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.50	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

différence=10%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	387	387	775
With Continuity Correction:	407	407	814

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.45	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.50	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

différence=5%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	1565	1565	3129
With Continuity Correction:	1604	1604	3209

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.47	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.50	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

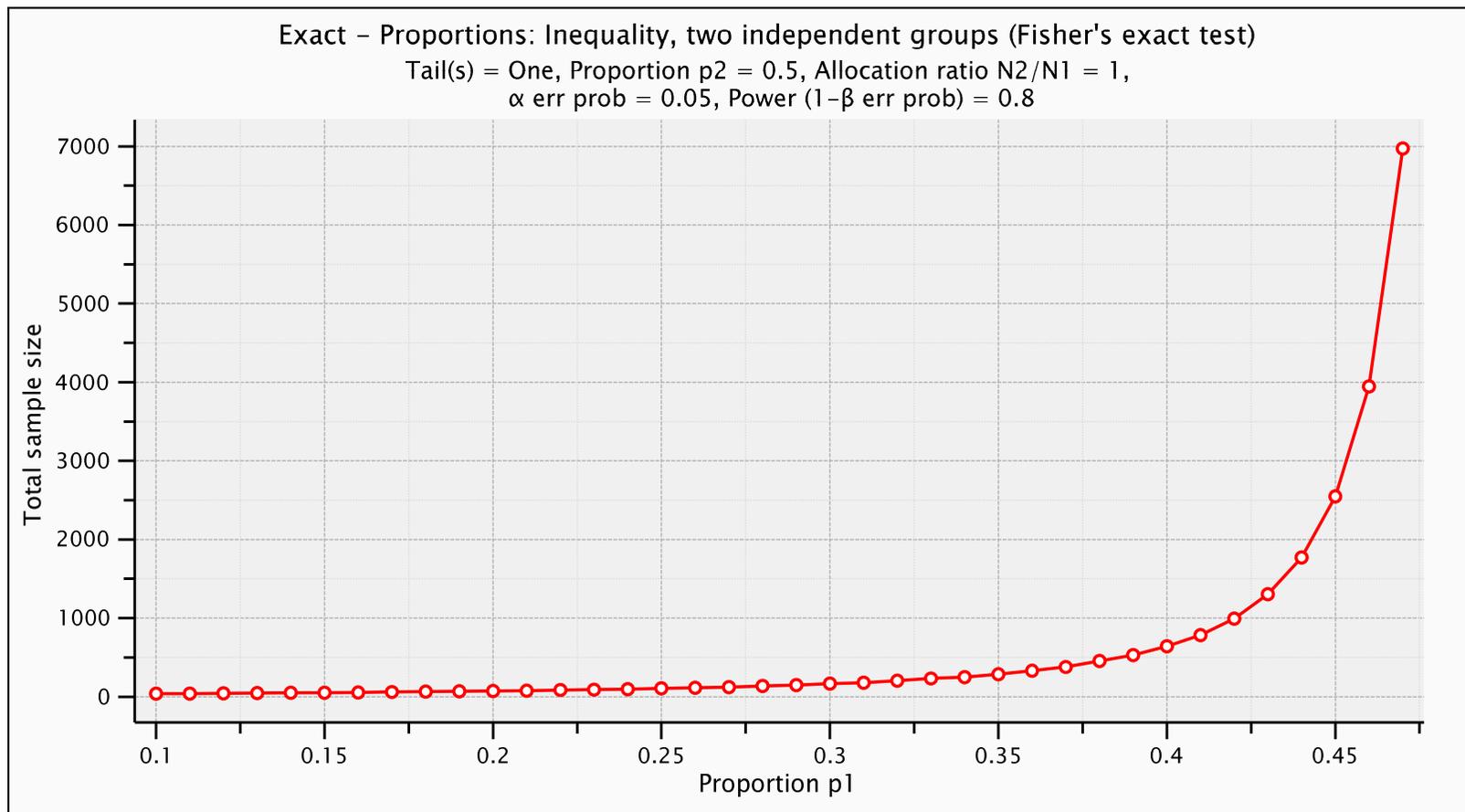
différence=3%

Compute

Sample Size Required

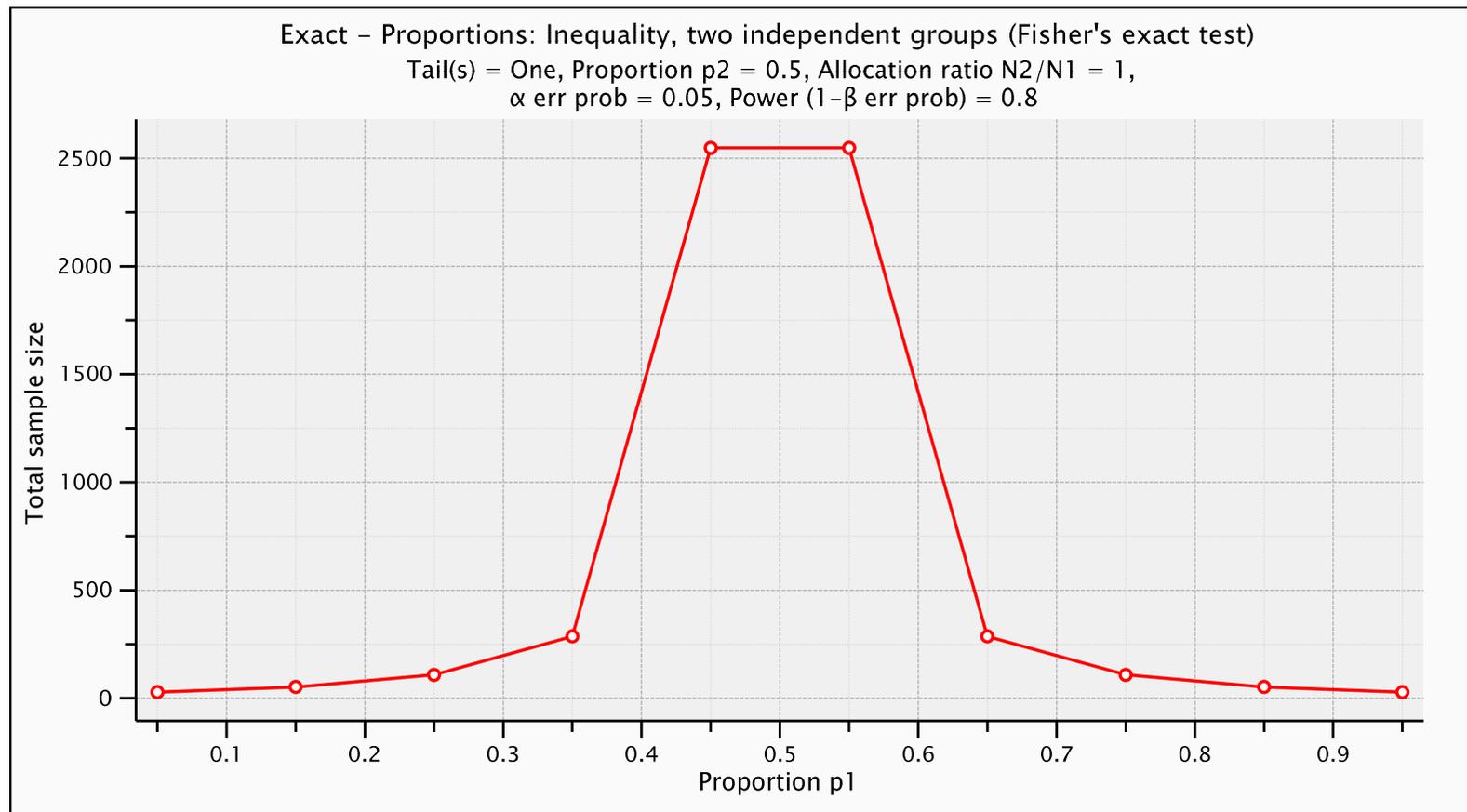
	Group 1	Group 2	Total
"Classical" Calculation:	4355	4355	8711
With Continuity Correction:	4422	4422	8844

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions <http://statpages.org/proppowr.html>



Taille totale des 2 échantillons en fonction des différentes valeurs de la proportion
- donc différences entre les deux échantillons de magnitude différente

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions <http://statpages.org/proppowr.html>



Taille totale des 2 échantillons en fonction des différentes valeurs du proportion
- donc différences entre les deux échantillons de magnitude différente

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions <http://statpages.org/proppowr.html>

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.40	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.50	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

50% vs. 40%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	387	387	775
With Continuity Correction:	407	407	814

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.30	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.40	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

40% vs. 30%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	356	356	712
With Continuity Correction:	376	376	751

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.20	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.30	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

30% vs. 20%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	293	293	586
With Continuity Correction:	313	313	626

Significance Level (alpha):	0.05	(Usually 0.05)
Power (% chance of detecting):	80	(Usually 80)
Group 1 Population Proportion:	.10	(Between 0.0 and 1.0)
Group 2 Population Proportion:	.20	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	1.0	(For equal samples, use 1.0)

20% vs. 10%

Compute

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	199	199	398
With Continuity Correction:	219	219	437

Même différence de 10%, mais pourcentages différents
Plus les pourcentages sont proche de 50% - plus les tailles sont grands

PLAN DE COLLECTION DES DONNÉES

Le calcul de taille de l'échantillon étudiée pour comparer deux proportions

<http://statpages.org/proppowr.html>

Relations entre le nombre des sujets et ...

Si on tiennent toutes les contraintes constantes (eg. Alpha, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux ...), **et on augmente** seulement la **puissance**, le **nombre des sujets va augmenter**

Si on tiennent toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, groupes de taille égaux / inégaux...), **et on augmente** seulement la **différence entre les pourcentages a observer**, le **nombre des sujets va diminuer**

Si on tiennent toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, groupes de taille égaux / inégaux ...), **et on diminue** seulement la **différence entre les pourcentages a observer**, le **nombre des sujets va augmenter**

Plus les pourcentages attendues qu'on compare sont proche a 50%, plus le nombre des sujets est grand

Plus les pourcentages attendues qu'on compare sont proche a 0%, plus le nombre des sujets est petit

Plus les pourcentages attendues qu'on compare sont proche a 100%, plus le nombre des sujets est petit

Des pourcentages attendues qu'on compare de distance égale a 50% (ex. 10% avec 90% ou 25% avec 75%) ont des nombres des sujets égaux

PLAN DE COLLECTION DES DONNÉES

Puissance – définition: probabilité d'observer une différence d'un certain magnitude si la différence est réelle

Erreur Beta (1 moins la puissance) – **définition:** probabilité ne pas observer une différence d'un certain magnitude si la différence est réelle

La puissance pour la différence entre deux proportions dépend de: nombre des sujets, différence entre les groupes, niveau de signification alpha, test unilatéral / bilatéral, groupes de taille égaux / inégaux

PLAN DE COLLECTION DES DONNÉES

Relations entre puissance et beta, nombre sujets, différence attendue, alpha

Si on tiennent toutes les contraintes constantes (eg. Alpha, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux ...), et on augmente seulement le **nombre des sujets, la **puissance va augmenter (et l'erreur beta va diminuer)****

Si on tiennent toutes les contraintes constantes (eg. Alpha, nombre des sujets, p uni/bidirectionnel, groupes de taille égaux / inégaux), et on augmente seulement la **différence entre les pourcentages a observer (critère de jugement – outcome/ endpoint – en anglais), la **puissance va augmenter (et l'erreur beta va diminuer)****

Si on tiennent toutes les contraintes constantes (eg. Alpha, nombre des sujets, p uni/bidirectionnel, groupes de taille égaux / inégaux ...), et on diminue seulement la **différence entre les pourcentages a observer (critère de jugement – outcome/ endpoint – en anglais), la **puissance va diminuer (et l'erreur beta va augmenter)****

Si on tiennent toutes les contraintes constantes (eg. nombre des sujets, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux ...), et on diminue seulement la **l'erreur alpha, la **puissance va diminuer (et l'erreur beta va augmenter)****

PLAN DE COLLECTION DES DONNÉES

Le **calcul de taille de l'échantillon étudiée** pour **comparer la moyenne** d'une caractéristique quantitative normale distribuée dans **deux groupes dépend** de:

- dimension de la différence entre les deux moyennes à observer (ex. 1)
- nature des données
- la valeur de la déviation standard = écart type (racine carrée de la variance) dans les deux groupes depuis des études antérieures
- aspects statistiques:
 - niveau de signification statistique alpha (ex. 5%)
 - puissance (80%)
 - utilisation du p bilatéral / unilatéral
 - groupes de dimension égaux / inégaux

<http://sampsizе.sourceforge.net/iface/s2.html#nm>

Le calcul de taille de l'échantillon étudiée pour comparer la moyenne d'une caractéristique quantitative normale distribuée dans deux groupes

<http://sampsizе.sourceforge.net/iface/s2.html#nm>

Meme difference de 20 mg/dl

Assumptions:

alpha = 5 (two-sided)
power = 80
m1 = 230
m2 = 210
sd1 = 26
sd2 = 33
n2/n1 = 1

Estimated sample size:

n1 = 35
n2 = 35

Puissance=80%

Assumptions:

alpha = 5 (two-sided)
power = 90
m1 = 230
m2 = 210
sd1 = 26
sd2 = 33
n2/n1 = 1

Estimated sample size:

n1 = 47
n2 = 47

Puissance=90%

Assumptions:

alpha = 5 (two-sided)
power = 95
m1 = 230
m2 = 210
sd1 = 26
sd2 = 33
n2/n1 = 1

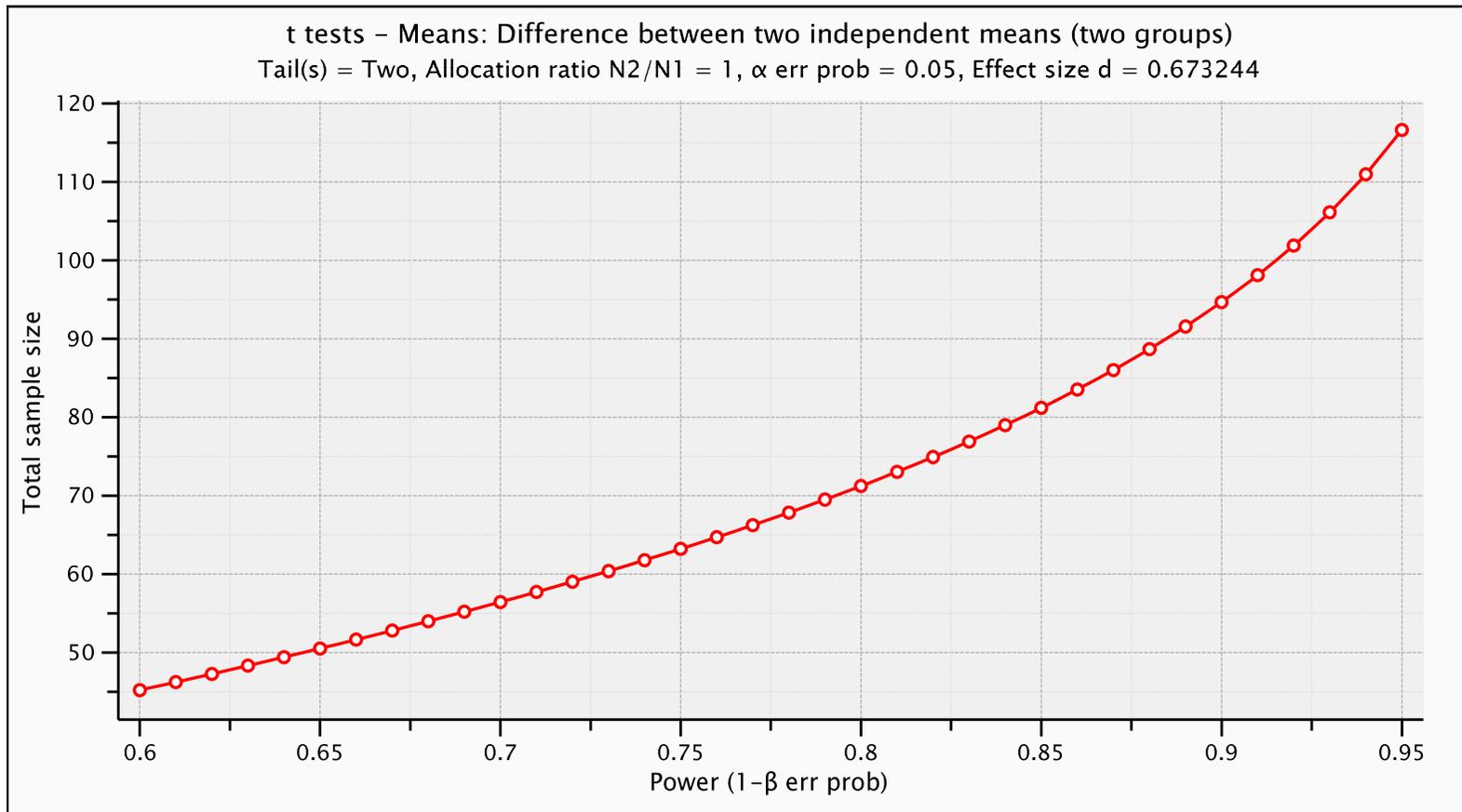
Estimated sample size:

n1 = 58
n2 = 58

Puissance=95%

PLAN DE COLLECTION DES DONNES

Le calcul de taille de l'échantillon étudiée pour comparer la moyenne d'une caractéristique quantitative normale distribuée dans deux groupes
<http://sampsizes.sourceforge.net/iface/s2.html#nm>



Taille totale des 2 échantillons en fonction de la puissance du test

Le calcul de taille de l'échantillon étudiée pour comparer la moyenne d'une caractéristique quantitative normale distribuée dans deux groupes

<http://sampsizе.sourceforge.net/iface/s2.html#nm>

Meme difference de 20 mg/dl

Assumptions:

alpha = 5 (two-sided)
power = 90
m1 = 230
m2 = 210
sd1 = 26
sd2 = 33
n2/n1 = 1

Estimated sample size:

n1 = 47
n2 = 47

230 vs. 210

Assumptions:

alpha = 5 (two-sided)
power = 90
m1 = 210
m2 = 190
sd1 = 26
sd2 = 33
n2/n1 = 1

Estimated sample size:

n1 = 47
n2 = 47

210 vs. 190

Assumptions:

alpha = 5 (two-sided)
power = 90
m1 = 190
m2 = 170
sd1 = 26
sd2 = 33
n2/n1 = 1

Estimated sample size:

n1 = 47
n2 = 47

190 vs. 170

PLAN DE COLLECTION DES DONNES

Le calcul de taille de l'échantillon étudiée pour comparer la moyenne d'une caractéristique quantitative normale distribuée dans deux groups

Assumptions:

```
alpha = 5 (two-sided)
power = 90
m1 = 230
m2 = 210
sd1 = 26
sd2 = 33
n2/n1 = 1
```

Estimated sample size:

```
n1 = 47
n2 = 47
```

différence=
230-210=20

Assumptions:

```
alpha = 5 (two-sided)
power = 90
m1 = 230
m2 = 220
sd1 = 26
sd2 = 33
n2/n1 = 1
```

Estimated sample size:

```
n1 = 186
n2 = 186
```

différence=
230-220=10

Assumptions:

```
alpha = 5 (two-sided)
power = 90
m1 = 230
m2 = 225
sd1 = 26
sd2 = 33
n2/n1 = 1
```

Estimated sample size:

```
n1 = 742
n2 = 742
```

différence=
230-225=5

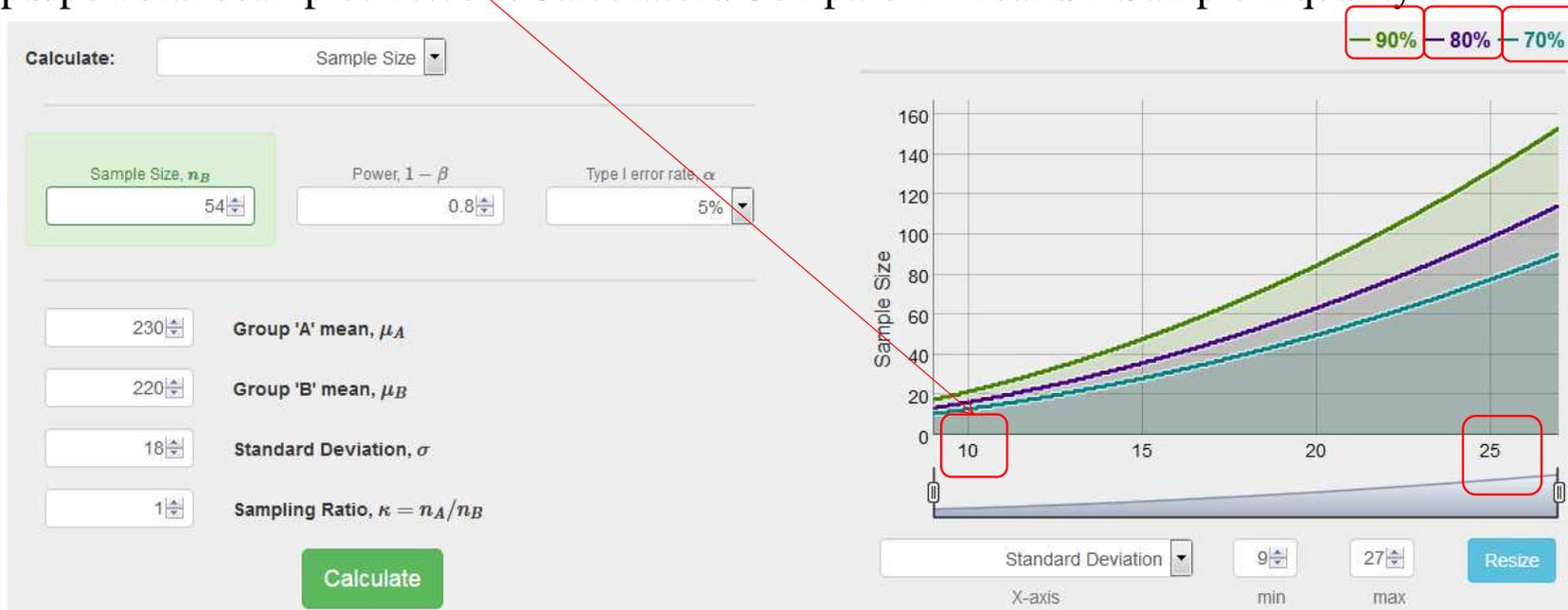
<http://sampsizes.sourceforge.net/iface/s2.html#nm>

PLAN DE COLLECTION DES DONNÉES

Le **calcul de taille de l'échantillon étudiée** pour comparer la **moyenne** d'une caractéristique quantitative normalement distribuée dans **deux groupes**

L'influence de la déviation standard sur la taille

<http://powerandsamplesize.com/Calculators/Compare-2-Means/2-Sample-Equality>



Déviatoin standard depuis 9 – 27 – sur l'axe horizontale, pour 3 puissances différentes

Le calcul de taille de l'échantillon étudiée pour comparer la moyenne d'une caractéristique quantitative normale distribuée dans deux groups

<http://sampsizе.sourceforge.net/iface/s2.html#nm>

Assumptions:

alpha = 5 (two-sided)
 power = 90
 m1 = 230
 m2 = 210
 sd1 = 26
 sd2 = 33
 n2/n1 = 1

Estimated sample size:

n1 = 47
 n2 = 47

Groupes egaux
n2/n1=1

Total=47+47=94

Assumptions:

alpha = 5 (two-sided)
 power = 90
 m1 = 230
 m2 = 210
 sd1 = 26
 sd2 = 33
 n2/n1 = 2

Estimated sample size:

n1 = 33
 n2 = 66

Groupes inegaux
n2/n1=2

Total=33+66=99

Assumptions:

alpha = 5 (two-sided)
 power = 90
 m1 = 230
 m2 = 210
 sd1 = 26
 sd2 = 33
 n2/n1 = 3

Estimated sample size:

n1 = 28
 n2 = 84

Groupes inegaux
n2/n1=3

Total=28+84=112

PLAN DE COLLECTION DES DONNES

Le **calcul de taille de l'échantillon étudiée** pour comparer la **moyenne** d'une caractéristique quantitative normale distribuée dans **deux groups**

Plus la taille des groups sont inégales (1 a 2, 1 a 3, 1 a 4), plus le nombre des sujets est grand

Plus la variance des données est grand (les déviations standards/ les écarts types – sont grand), plus le nombre des sujets est grand

PLAN DE COLLECTION DES DONNÉES

Le **calcul de taille de l'échantillon étudiée** pour comparer la **moyenne** d'une caractéristique quantitative normale distribuée dans **deux groups**

Relations entre le nombre des sujets et ...

Si on tiennent toutes les contraintes constantes (eg. Alpha, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux, les déviations standards/ les écarts types/les variances ...), **et on augmente** seulement la **puissance**, le **nombre des sujets va augmenter**

Si on tiennent toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, groupes de taille égaux / inégaux, les déviations standards/ les écarts types/les variances ...), **et on augmente** seulement la **différence entre les moyennes a observer**, le **nombre des sujets va diminuer**

Si on tiennent toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, groupes de taille égaux / inégaux, les déviations standards/ les écarts types/les variances . . .), **et on diminue** seulement la **différence entre les moyennes a observer**, le **nombre des sujets va augmenter**

Les valeurs des moyennes attendues ne compte pas (n'ont pas une influence sur la taille), seulement la différence entre les moyennes influence le nombre des sujets.

PLAN DE COLLECTION DES DONNES

Le **calcul de taille de l'échantillon étudiée** pour comparer la **moyenne** d'une caractéristique quantitative normale distribuée dans **deux groups**

Relations entre le nombre des sujets et ...

Si on tient toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux, ...), **et on augmente seulement la variabilité des donnes – donc: les déviations standards/ les écarts types/les variances, le nombre des sujets va diminuer**

Si on tient toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux, ...), **et on diminue seulement la variabilité des donnes – donc: les déviations standards/ les écarts types/les variances, le nombre des sujets va augmenter**

On peut diminuer la variabilité, par exemple en utilisant une technique de mesure plus précise, moins variable

Si on tient toutes les contraintes constantes (eg. Alpha, puissance, p uni/bidirectionnel, différence a obs...), **et on augmente seulement l'inegalite entre les tailles des groups** (1 a 2, 1 a 3, 1 a 4), **le nombre des sujets va augmenter**

PLAN DE COLLECTION DES DONNES

La **puissance d'un étude** pour comparer la **moyenne** d'une caractéristique quantitative normale distribuée dans **deux groups**

Relations entre la **puissance** et la différence entre deux moyennes et le nombre des sujets, beta, différence entre les moyennes, erreur alpha, la variance, la déviation standard

Si on tiennent toutes les contraintes constantes (eg. Alpha, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux ...), **et on augmente** seulement **le nombre des sujets**, la **puissance** va **augmenter** (et l'erreur beta va **diminuer**)

Si on tiennent toutes les contraintes constantes (eg. Alpha, nombre des sujets, p uni/bidirectionnel, groupes de taille égaux / inégaux), **et on augmente** seulement la **différence entre les moyennes a observer** (critère de jugement – outcome/ endpoint – en anglais), la **puissance** va **augmenter** (et l'erreur beta va **diminuer**)

Si on tiennent toutes les contraintes constantes (eg. Alpha, nombre des sujets, p uni/bidirectionnel, groupes de taille égaux / inégaux ...), **et on diminue** seulement la **différence entre les moyennes a observer** (critère de jugement – outcome/ endpoint – en anglais), la **puissance** va **diminuer** (et l'erreur beta va **augmenter**)

PLAN DE COLLECTION DES DONNES

La **puissance d'un étude** pour comparer la **moyenne** d'une caractéristique quantitative normale distribuée dans **deux groups**

Relations entre la **puissance** et la différence entre deux moyennes et le nombre des sujets, **beta**, différence entre les moyennes, erreur **alpha**, la variance, la déviation standard

Si on tiennent toutes les contraintes constantes (eg. nombre des sujets, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux ...), **et on diminue** seulement la **l'erreur alpha**, la **puissance** va **diminuer** (et **l'erreur beta** va **augmenter**)

Si on tiennent toutes les contraintes constantes (eg. Alpha, nombre des sujets, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux, ...), **et on augmente** seulement la **variabilité des donnes** – **donc: les déviations standards/ les écarts types/les variances**, la **puissance** va **diminuer** (et **l'erreur beta** va **augmenter**)

Si on tiennent toutes les contraintes constantes (eg. Alpha, nombre des sujets, p uni/bidirectionnel, différence a obs, groupes de taille égaux / inégaux, ...), **et on diminue** seulement la **variabilité des donnes** – **donc: les déviations standards/ les écarts types/les variances**, la **puissance** va **augmenter** (et **l'erreur beta** va **diminuer**)

PLAN DE COLLECTION DES DONNÉES

Concernant la puissance statistique:

- Dépends
 - de nombre de patients analysées,
 - du nombre des sujets inclus
 - du nombre de sujets qui ont subi l' évènement étudié
 - de la différence sur le critère de jugement entre les groupes de l' étude
- Un nombre de sujets nécessaire est calculé pour avoir une puissance donnée
- Est reflétée par la largeur de l'intervalle de confiance
 - Grande puissance – intervalle étroit
 - Petite puissance – intervalle large

PLAN DE COLLECTION DES DONNÉES

- Définition claire des variables
- Etablir les échelles de mesure
- Définir les critères de normalité et les méthodes diagnostiques utilisées
- Définir le type de collection des données en fonction du:
 - Type et des objectifs de l'étude
 - Modalités des choix des sujets
 - Disponibilité de temps
 - Ressources financières et humaines
 - Procédure utilisée
 - Accès aux données
- Identification des potentiels facteurs d'erreur
- Définition du plan d'analyse statistique

PLAN DE COLLECTION DES DONNES

- **La réalisation correcte de la base des données qui contiendra les variables d'étude:**
- La première ligne contiennent les noms des variables, avec l'unité de mesure.
 - erreur: avoir deux lignes pour cela ou commencer les données depuis ≥ 2 ème ligne
- Si on a plus des groupes a comparer
 - on met toutes les données sur la même feuille de calcul,
 - l'une dans la continuation de l'autre,
 - on identifie l'appartenance au group par une colonne (variable) de regroupage.

PLAN DE COLLECTION DES DONNÉES

- **La réalisation correcte de la base des données qui contiendra les variables d' étude:**
- Si on a des variables calculables a partir d'autre variables on utilisent des fonctions des calcul tabellaire
 - (ex. l'indice de masse corporelle – en fonction de l'hauteur et du poids)
- On met des méthodes d'identification des données: n, feuille d'observation/nom prénom/ code d'identification.
- S'il y a des données manquantes on laisse ces cellules libres
- Après la collection des données on doit vérifier les données pour des valeurs aberrantes/ extrêmes.
 - Si on les identifient on doit les modifier avec des valeurs réels, et non avec des suppositions
 - Si on ne trouve pas les valeur originales c'est mieux de les considérer comme des valeur inconnues que d'inventer des valeurs
 - On doit pas effacer des valeurs extrêmes comme une règle

Préparation de l'étude

- Si c'est une étude thérapeutique: présentation de:
 - les interventions
 - la méthode aveugle
 - l'allocation
 - la randomisation
 - des traitements concomitants admis/non
 - la compliance au traitement

Préparation de l'étude

- Si c'est une étude thérapeutique: présentation de:
 - Evaluation de l'efficacité
 - Critères principales/secondaires
 - Définition des « endpoint » principales/secondaires
 - La sécurité des patients
 - Effets secondaires
 - Définitions
 - Types:
 - Pas grave
 - grave
 - Procédures pour la retrait des patients
 - Les critères de retrait

Préparation de l'étude

- Des considérations statistiques
- Le plan des analyses statistiques
 - Description des données
 - Méthodes statistiques
 - Population à analyser (intention à traiter)
- Le calcul de la taille de l'échantillon
- Procédures pour des données qui manquent
- Mesure de la compliance
- Des traitements concomitantes
- Des comparaisons multiples
- Analyse des sous-groupes

Préparation de l'étude

- Les procédures d'étude
- La programmation des visites
 - Avant l'étude
 - Au début de l'étude
 - Description de chaque visite
- Considérations pratiques:
 - Moyennes financières, le sponsor
 - Stabilité et qualité du personnel et de l'équipement pendant l'étude
 - Considérations éthiques
 - Consentement éclairée
 - Confidentialité
 - Audit

Préparation de l'étude

L'étape de préparation de l'étude doit se finaliser dans une manière obligatoire avec un **protocole écrit**, contenant les éléments antérieures.

Validation par l'équipe

Aucune modification après la validation

Exécution de l'étude

- Collection et analyse des données
- Présentation des résultats préliminaire
- Analyse et interprétation des résultats
 - Hypothèse confirmée ou rejetée?
 - Analyse numérique suffisante?
 - Quelle est la signifiante des résultats?

Présentation et rédaction des résultats

- **Nouveau documentation bibliographique** pour éviter de perdre des résultats importants publiés pendant le déroulement de l'étude → chapitre "Discussions".
- Présentation des résultats en utilisant des **paramètres statistiques** qui permettent de rapporter les résultats aux résultats des études similaires, et la reproduction de notre étude.

Les Etapes d'une Recherche

1. Préparation de l'étude
2. Le protocole de recherche écrit
3. Collection et analyse des données
4. L'analyse des résultats préliminaires
5. Analyse et interprétation des résultats
6. Rédaction de l'études



Exemples des questions pour l'examen

Q1) Quelles types de présentation vous suggérez pour une étude qui a mesure la longueur du nez, et ils ont observé que les données sont normalement distribuées?

- a) graphique des moyennes
- b) graphique camembert
- c) moyenne et déviation standard
- d) médiane et déviation standard
- e) médiane et intervalle interquartile

Réponse: a, c

Q2) Quelles types de présentation vous suggérez pour une étude qui a mesure la longueur du mandibule, et ils ont observé que les données ne sont pas normalement distribuées?

- a) graphique des moyennes
- b) graphique boîte à moustaches
- c) moyenne et déviation standard
- d) médiane et déviation standard
- e) médiane et intervalle interquartile

Réponse: b, e

Exemples des questions pour l'examen

Q3) Quelles types de présentation vous suggérez pour une étude qui a observée le présence d'un cancer de lèvre?

- a) graphique des barres
- b) graphique camembert
- c) tableau des fréquences
- d) graphique des colonnes
- e) médiane et intervalle interquartile

Réponse: a, b, c, d

Q4) * Quelle type de présentation vous suggérez pour une étude qui a mesure l' évolution de la tension artérielle systolique pour 10 jours?

- a) graphique ligne
- b) graphique de survie
- c) graphique camembert
- d) tableau de contingence
- e) tableau des fréquences

Réponse: a

Exemples des questions pour l'examen

Q5) Quelles types de présentation vous suggérez pour une étude qui a observée le temps entre un intervention chirurgicale pour éliminer une cancer du lèvre et le décès?

- a) graphique des barres
- b) graphique camembert
- c) la probabilité de survie a 12 mois
- d) graphique du survie
- e) médiane et intervalle interquartile

Réponse: c, d, e

Q6) Quelles types de présentation vous suggérez pour une étude qui a évaluée la relation entre la consommation d'alcool (oui/non) et le cancer de lèvre (oui/non)?

- a) graphique ligne
- b) graphique des colonnes
- c) graphique camembert
- d) tableau de contingence
- e) graphique nouage des points (scatter)

Réponse: b, d

Exemples des questions pour l'examen

Q7) * Quelle type de présentation vous suggérez pour une étude qui a évaluée la relation entre la longueur du nez et la longueur du mandibule?

- a) graphique ligne
- b) graphique des colonnes
- c) graphique camembert
- d) tableau de survie
- e) graphique nouage des points (scatter)

Réponse: e

Q8) Quelles types de présentation vous suggérez pour une étude qui a évaluée la relation entre la consommation d'alcool (oui/non) et le poids du patients (en kg), sachant que le poids est normale distribuée?

- a) graphique des moyennes (avec déviation standard)
- b) graphique colonnes avec barre d' erreur
- c) moyenne et déviation standard par groupes
- d) graphique ligne
- e) médiane et intervalle interquartile

Réponse: a, b, c

Exemples des questions pour l'examen

Q11) Quelles types de présentation vous suggérez pour une étude qui a évaluée la relation entre la consommation d'alcool (oui/non) et le taux de l'insuline dans le sang, sachant que l'insuline n'est pas normale distribuée?

- a) graphique des moyennes (avec déviation standard)
- b) graphique colonnes avec barre d'erreur
- c) graphique boîte a moustaches
- d) graphique ligne
- e) médiane et intervalle interquartile

Réponse: c, e

Q12) Vous avez trouvé deux études sur Pubmed qui ont été réalisée pour évaluer la relation entre la consommation exagérée du the comme facteur de risque pour l'apparition de l'ostéoporose. Ils ont obtenu le suivant OR avec l'intervalle de confiance : étude A : 2 (95% CI 1,15-41), et étude B : 2,9 (95% CI, 2,4 – 15). On considère que les valeurs supérieure a 2 sont cliniquement importants, et les valeurs inferieures a 1,3 – que ne sont pas cliniquement importants. Lesquelles des suivantes affirmations sont correctes ?

- a) l'étude A est moins précis que l'étude B
- b) l'étude A est un nombre plus petit des patients que l'étude B
- c) l'étude A est un nombre plus grand des patients que l'étude B
- d) l'intervalle de confiance d'étude A n'est pas claire du point de vue cliniquement
- e) l'intervalle de confiance d'étude B est cliniquement important

Réponse: a, b, d, e

Exemples des questions pour l'examen

Q13) Vous avez trouvé deux études sur Pubmed qui ont été réalisées pour évaluer la différence de temps de fonctionnement entre deux prothèses du genou sans des problèmes. Ils ont obtenu les suivantes moyennes des différences avec l'intervalle de confiance : étude A : 12 mois (95% CI 10-14), et étude B : 10 mois (95% CI, 9,5 – 10,5). On considère que les valeurs supérieures à 12 mois sont cliniquement importants, et les valeurs inférieures à 12 mois – que ne sont pas cliniquement importants. Lesquelles des suivantes affirmations sont correctes ?

- a) l'étude B est plus puissante que l'étude A
- b) l'étude A est un nombre plus petit des patients que l'étude B
- c) l'étude A est un nombre plus grand des patients que l'étude B
- d) l'intervalle de confiance d'étude A n'est pas clair du point de vue clinique
- e) l'intervalle de confiance d'étude B n'est pas cliniquement important

Réponse: a, b, d, e

Exemples des questions pour l'examen

Q14) Une étude a été réalisée sur un seul échantillon des 5000 médecins tirée au hasard depuis toutes les médecins d'un pays pour évaluer la relation entre la consommation régulière d'alcool (oui/non) dans les dernières trois années et le poids du patients (en kg). Les chercheurs ont mesurée le poids médecins et ont utilise une questionnaire pour trouver la consommation régulière d'alcool dans les dernières trois années. Lesquelles des suivantes caractéristiques sont juste pour cet étude?

- a) domaine de la recherche clinique - description d'un phénomène de la santé
- b) domaine de la recherche clinique - évaluations des facteurs de risque ou pronostiques
- c) domaine de la recherche clinique - évaluation d'un test diagnostique
- d) domaine de la recherche clinique - évaluation d'un abord thérapeutique
- e) étude descriptive
- f) étude analytique
- g) étude observationnelle
- h) étude expérimentale
- i) collection des données transversale
- j) collection des données longitudinale
- k) collection des données rétrospective
- l) collection des données prospective
- m) collection des données cas - témoins
- n) collection des données exposée - non exposée
- o) collection des données échantillon représentative
- p) collection des données exhaustive
- q) collection des données par échantillonnage

Réponse: b, f, g, j, k, o, q

Exemples des questions pour l'examen

Q15) Une étude a été réalisée sur un seul échantillon des 5000 médecins tirée au hasard depuis toutes les médecins d'un pays pour évaluer l'association entre la longueur du nez et la longueur du mandibule. Les données ont été mesurées sur chaque médecin dans quelques jours. Lesquelles des suivantes caractéristiques sont justes pour cette étude?

- a) domaine de la recherche clinique - description d'un phénomène de la santé
- b) domaine de la recherche clinique - évaluations des facteurs de risque ou pronostiques
- c) domaine de la recherche clinique - évaluation d'un test diagnostique
- d) domaine de la recherche clinique - évaluation d'un abord thérapeutique
- e) étude descriptive
- f) étude analytique
- g) étude observationnelle
- h) étude expérimentale
- i) collection des données transversale
- j) collection des données longitudinale
- k) collection des données rétrospective
- l) collection des données prospective
- m) collection des données cas - témoins
- n) collection des données exposée - non exposée
- o) collection des données échantillon représentative
- p) collection des données exhaustive
- q) collection des données par échantillonnage

Réponse: b, f, g, i, l, o, q

Exemples des questions pour l'examen

Q16) Une étude a été réalisée sur un seul échantillon des 5000 médecins tirée au hasard depuis toutes les médecins d'un pays pour évaluer la présence de l'obésité. Les données de leur dernière mesure du poids et de l'hauteur nécessaire pour l'évaluation de l'obésité ont été trouvées dans la base des données du médecin de travail. Lesquelles des suivantes caractéristiques sont juste pour cet étude?

- a) domaine de la recherche clinique - description d'un phénomène de la santé
- b) domaine de la recherche clinique - évaluations des facteurs de risque ou pronostiques
- c) domaine de la recherche clinique - évaluation d'un test diagnostique
- d) domaine de la recherche clinique - évaluation d'un abord thérapeutique
- e) étude descriptive
- f) étude analytique
- g) étude observationnelle
- h) étude expérimentale
- i) collection des données transversale
- j) collection des données longitudinale
- k) collection des données rétrospective
- l) collection des données prospective
- m) collection des données cas - témoins
- n) collection des données exposée - non exposée
- o) collection des données échantillon représentative
- p) collection des données exhaustive
- q) collection des données par échantillonnage

Réponse: a, e, g, i, k, o, q

Exemples des questions pour l'examen

Q17) Une étude a été réalisée sur un seul échantillon des 100 sujets pour évaluer si le diabète influence l'évolution de la parodontite. Ils ont identifié 50 sujets avec du diabète sucré, avec parodontite, et 50 sujets sans diabète, avec parodontite et ils ont observé l'évolution de la parodontite dans plusieurs visites médicales, après le début de l'étude. Lesquelles des suivantes caractéristiques sont justes pour cette étude?

- a) domaine de la recherche clinique - description d'un phénomène de la santé
- b) domaine de la recherche clinique - évaluations des facteurs de risque ou pronostiques
- c) domaine de la recherche clinique - évaluation d'un test diagnostique
- d) domaine de la recherche clinique - évaluation d'un abord thérapeutique
- e) étude descriptive
- f) étude analytique
- g) étude observationnelle
- h) étude expérimentale
- i) collection des données transversale
- j) collection des données longitudinale
- k) collection des données rétrospective
- l) collection des données prospective
- m) collection des données cas - témoins
- n) collection des données exposée - non exposée
- o) collection des données échantillon représentative
- p) collection des données exhaustive
- q) collection des données par échantillonnage

Réponse: **b, f, g, j, l, n, q**

Exemples des questions pour l'examen

Q17) Quelles affirmations concernant les types des variables sont correctes :

- a) nombre des patients d'un médecin – variable quantitative discrète
- b) nombre des opérations d'un chirurgienne – variable quantitative continue
- c) la tension artérielle systolique – variable quantitative discrète
- d) le temps de bossage des dents – variable quantitative continue
- e) la concentration du créatinine dans le sang – variable quantitative discrète

Réponse: a, d

Q18) Quelles affirmations concernant les types des variables sont correctes :

- a) nombre des dents dans la bouche – variable quantitative discrète
- b) nombre des enfants dans une famille – variable quantitative continue
- c) nombre des dents caries – variable quantitative discrète
- d) la fréquence de brossage par jour – variable quantitative continue
- e) la concentration du cholestérol dans le sang – variable quantitative continue

Réponse: a, c, e

Exemples des questions pour l'examen

Q19) Quelles affirmations concernant les facteurs qui influence la puissance d'un étude sont correctes :

- a) Plus le nombre des sujets est grand plus la puissance est forte
- b) Plus le nombre des sujets est grand plus l'erreur Beta est petite
- c) Plus la différence entre les moyennes attendues (critère de jugement – outcome/ endpoint – en anglais) qu'on compare est petite, plus la puissance diminue
- d) Plus alpha est petit – la puissance diminue
- e) Plus beta est grand – la puissance diminue
- f) Plus la variance (le carré de la déviation standard) / variation des données est grand (les déviations standards/ les écarts types – sont grand), plus la puissance est réduite

Réponse: a, b, c, e, f

Q20) Quelles affirmations concernant les facteurs qui influence la puissance d'un étude sont correctes :

- a) de nombre de patients analysées,
- b) du nombre des sujets inclus
- c) du nombre de sujets qui ont subi l'évènement étudié
- d) de la différence sur le critère de jugement entre les groupes de l'étude
- e) du niveau de signifiante

Réponse: a, b, c, d, e

Exemples des questions pour l'examen

Q21) Quelles types de méthodes vous suggérez pour une étude qui a évaluée la relation entre la consommation d'alcool (oui/non) et le poids du patients (en kg), sachant que le poids est normale distribuée?

- a) tests statistiques paramétriques
- b) tests statistiques comme le test t
- c) moyenne et déviation standard par groupes
- d) graphique ligne
- e) médiane et intervalle interquartile

Réponse: a, b, c

Q22) Quelles méthodes vous suggérez pour une étude qui a évaluée la relation entre la consommation d'alcool (oui/non) et le taux de l'insuline dans le sang, sachant que l'insuline n'est pas normale distribuée?

- a) tests statistiques non paramétriques
- b) tests statistiques comme le test Mann Whitney U
- c) graphique boite a moustaches
- d) tests statistiques comme le test t
- e) tests statistiques comme le test Wilcoxon rank sum test

Réponse: a, b, c, e

Exemples des questions pour l'examen

Q23) On a deux études qui évaluent la différence entre la réduction du taux de cholestérol entre deux médicaments. L'étude A a trouvé une différence moyenne de 5 mg/dL, avec une déviation standard de 8mg/dL, et l'étude B a trouvé une différence moyenne de 3 mg/dL, avec une déviation standard de 6mg/dL. Quelles affirmations sont correctes considérant pour chaque réponse que toutes les autres facteurs qui influencent le nombre des sujets, sont identiques entre les deux études, sauf ceux précisés dans la réponse:

- a) Si on regarde seulement l'écart type de l'étude B, par rapport à l'étude A, l'étude B demande moins des sujets
- b) Si on regarde seulement la différence moyenne de l'étude B, par rapport à l'étude A, l'étude A demande plus des sujets
- c) L'écart type de l'étude B est plus grand que la déviation standard de l'étude A
- d) Si on veut une plus grande puissance statistique pour l'étude B, on doit augmenter le nombre des sujets
- e) La variance de l'étude A est plus grande que celle de l'étude B

Réponse: a, c, d, e

Fin



**Merci pour
votre attention!**

