

**Etudes descriptives.
Modélisation des phénomènes médicaux (Analyse
de régression)**

Objectifs

- Recherche (étude) descriptive
- Analyse de Régression:
 - Régression logistique simple
 - Régression logistique multiple

Recherche descriptive

- Les **études descriptives** sont des études observationnelles qui décrivent:
 - les caractéristiques d'un phénomène,
 - les caractéristiques d'un sujet (patient)
 - les caractéristiques d'une population étudiée;
- est utilisé lorsqu'il y a peu d'informations sur phénomène
- une recherche préliminaire;

Type de recherche

Type de recherche		Question	Direction de recherche
Descriptive		Que se passe t-il ? Nombre de cas?	description des certains phénomènes de santé (ex. une nouvelle maladie infectieuse)
Analytique	Explicative	Quelles sont les facteurs qui expliquent le phénomène de santé/résultat clinique d'intérêt ?	Construire une théorie (modèle) en testant des hypothèses
	Predictive	Que se passera t-il si... ?	Tester et vérifier des hypothèses

- les **études descriptives** essaient **juste de comprendre le phénomène** étudié sans chercher à expliquer pourquoi ce phénomène se produit
- Les études analytiques font des comparaisons, essayant d'expliquer le phénomène

Types d'études (recherches) descriptives

- Étude de cas
- Série de cas
- Études transversales (études de prévalence)
- Etudes écologiques (et corrélationnelles)

Étude de cas (Cas rapporté)

- décrit l'expérience d'un seul patient ou d'un groupe de patients avec un diagnostic similaire.
- décrit comment un clinicien identifie une caractéristique inhabituel d'une maladie ou des antécédents d'un patient
- peut être le premier indice pour identifier des nouvelles maladies ou effets néfastes d'une exposition
- un rapport de cas peut nécessiter des investigations de conception supplémentaires

Étape suivante : série de cas

Exemple

Case Reports > Case Rep Med. 2016;2016:5750710. doi: 10.1155/2016/5750710.

Epub 2016 Dec 26.

Fetus in the Bladder: Rare Complication of Vesicouterine Fistula

Vladimir Lesovoy ¹, Yuri Parashchuk ², Dmytro Shchukin ¹, Roman Safonov ³, Karyna Didenkova ⁴, Maria Lisova ², Inessa Safonova ⁵

FULL TEXT LINKS



ACTIONS

“ Cite

Abstract

The work presents a rare case of spontaneous migration of an 11-week fetus from the uterine cavity into the urinary bladder cavity through the long-standing vesicouterine fistula.

Conclusions

The onset of pregnancy is possible in the presence of the long-term vesicouterine fistula.

The migration of the fetus from the uterine cavity into the bladder presents one of the complications of vesicouterine fistulas.

Lien vers l'article: <https://pubmed.ncbi.nlm.nih.gov/28105051/>

Série de cas

- groupe de patients semblables entre eux (ayant une caractéristique commune d'intérêt) = les « cas »
- il n'y a pas de groupe témoin ou une autre groupe de patients impliqué dans l'étude
- synthèse d'observations cliniques de cas individuels survenant dans une période relativement courte de temps.

Quand il est d'intérêt d'utiliser les séries de cas?

- on observe un phénomène inhabituel (une nouvelle maladie)
- peu de connaissances sur un domaine médical (nouvelle pratique)
- on veut des données/preuves « scientifiques » (en regroupant des observations cliniques similaires) pour formaliser un phénomène médical « intuitif »

Série de cas : **Avantages et désavantages**

- **Avantages :**

- ✓ rapide, peu coûteux,

- **Désavantages:**

- ✓ Pas de groupe de comparaison
 - Pas de mise en évidence d'un lien de causalité/lien de dépendance
 - Impossible de savoir si les facteurs observés sont plus fréquents parmi les cas que dans la population (ne peut pas être utilisé pour tester l'association)
- ✓ Provient d'observations cliniques : ne représente pas la population à risque
- ✓ Ne représente pas l'ampleur/la magnitude du phénomène

Étape suivante : étude de prévalence

Exemple

> Nature. 2020 Mar;579(7798):270-273. doi: 10.1038/s41586-020-2012-7. Epub 2020 Feb 3.

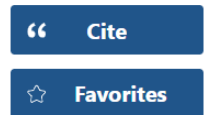
A pneumonia outbreak associated with a new coronavirus of probable bat origin

Peng Zhou ^{#1}, Xing-Lou Yang ^{#1}, Xian-Guang Wang ^{#2}, Ben Hu ¹, Lei Zhang ¹, Wei Zhang ¹, Hao-Rui Si ^{1 3}, Yan Zhu ¹, Bei Li ¹, Chao-Lin Huang ², Hui-Dong Chen ², Jing Chen ^{1 3}, Yun Luo ^{1 3}, Hua Guo ^{1 3}, Ren-Di Jiang ^{1 3}, Mei-Qin Liu ^{1 3}, Ying Chen ^{1 3}, Xu-Rui Shen ^{1 3}, Xi Wang ^{1 3}, Xiao-Shuang Zheng ^{1 3}, Kai Zhao ^{1 3}, Quan-Jiao Chen ¹, Fei Deng ¹, Lin-Lin Liu ⁴, Bing Yan ¹, Fa-Xian Zhan ⁴, Yan-Yi Wang ¹, Geng-Fu Xiao ¹, Zheng-Li Shi ⁵

FULL TEXT LINKS



ACTIONS



• COVID-19 - Premier rapport d'une série de patients

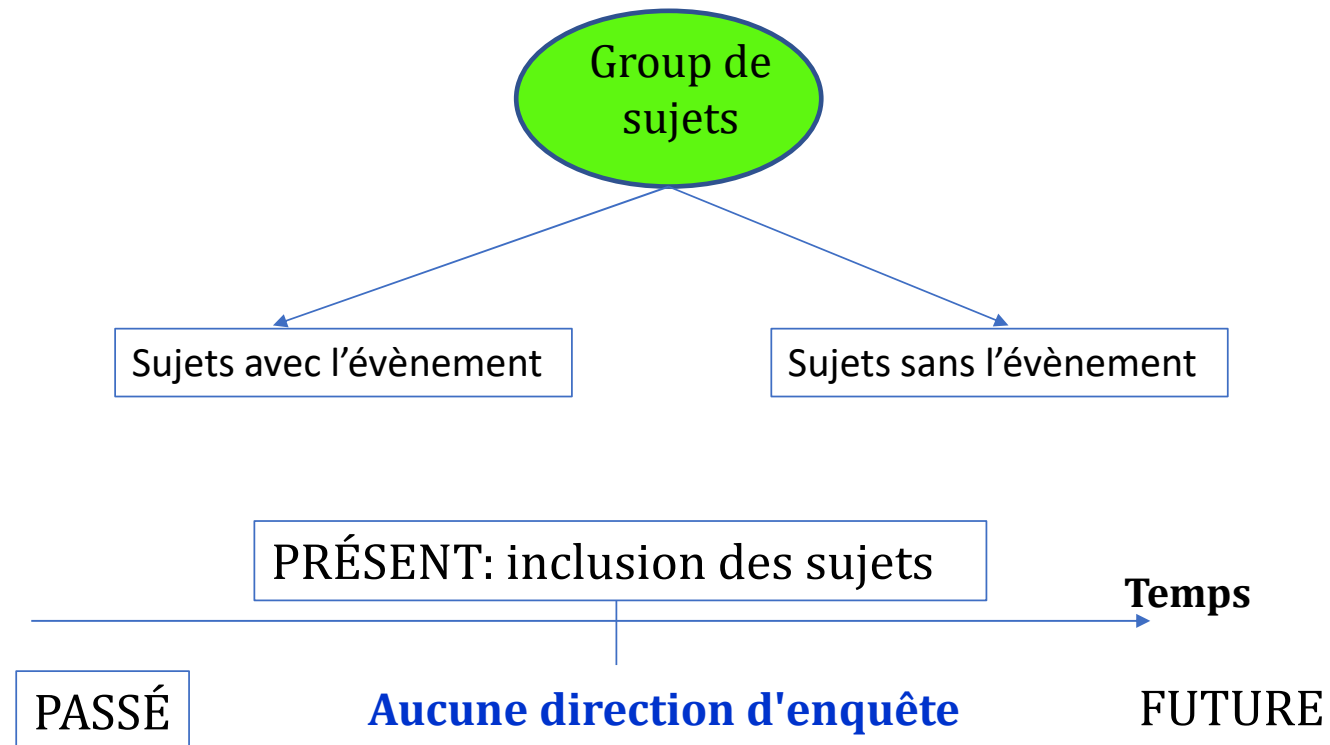
Abstract

Since the outbreak of severe acute respiratory syndrome (SARS) 18 years ago, a large number of SARS-related coronaviruses (SARSr-CoVs) have been discovered in their natural reservoir host, bats¹⁻⁴. Previous studies have shown that some bat SARSr-CoVs have the potential to infect humans⁵⁻⁷. [Here we report the identification and characterization of a new coronavirus \(2019-nCoV\)](#), which caused an epidemic of acute respiratory syndrome in humans in Wuhan, China. The epidemic, which started on 12 December 2019, had caused 2,794 laboratory-confirmed infections including 80 deaths by 26 January 2020. [Full-length genome sequences were obtained from five patients at an early stage of the outbreak](#). The sequences are almost identical and share 79.6% sequence identity to SARS-CoV. Furthermore, we show that 2019-nCoV is [96% identical at the whole-genome level to a bat coronavirus](#). Pairwise protein sequence analysis of seven conserved non-structural proteins domains show that this virus belongs to the species of SARSr-CoV. In addition, 2019-nCoV virus isolated from the bronchoalveolar lavage fluid of a critically ill patient could be neutralized by sera from several patients. Notably, we confirmed that 2019-nCoV uses the same cell entry receptor-angiotensin converting enzyme II (ACE2)-as SARS-CoV. (lien vers l'article: <https://www.nature.com/articles/s41586-020-2012-7>)

Étude transversale (étude de prévalence): définition

- engl. *Cross-sectional study*
- Mesurer la **fréquence** d'un phénomène de sante:
 - maladie
 - facteur d'exposition
 - exposition à un traitement
- Enquête sur les sujets/patients présents au moment du recueil
- Une seule observation par sujet/patient = « **photographie** » de l'état de la population étudiée

Schéma d'une étude transversale: Que se passe-t-il en ce moment?



Exemple

Acta Ophthalmol. 2020 Dec;98(8):e951-e959. doi: 10.1111/aos.14472. Epub 2020 May 18.

Ocular manifestations and clinical characteristics of 535 cases of COVID-19 in Wuhan, China: a cross-sectional study

Liwen Chen¹, Chaohua Deng¹, Xuhui Chen¹, Xian Zhang¹, Bo Chen¹, Huimin Yu¹, Yuanjun Qin¹, Ke Xiao¹, Hong Zhang¹, Xufang Sun¹

FULL TEXT LINKS

WILEY Full Text Article

PMC FREE Full text

ACTIONS

“ Cite

- **Abstract**
- **Purpose:** To investigate the ocular manifestations and clinical characteristics of COVID-19 patients caused by SARS-CoV-2 in Wuhan, China.
- **Methods:** A total of 535 COVID-19 patients were recruited at Mobile Cabin Hospital and Tongji Hospital. Information on demographic characteristics, exposure history, ocular symptoms, eye drop medication, eye protections, chronic eye diseases, systemic concomitant symptoms, radiologic findings and SARS-CoV-2 detection in nasopharyngeal swabs by real-time PCR was collected from questionnaires and electronic medical records.
- **Results:** Of 535 patients, 27 patients (5.0%) presented with conjunctival congestion and 4 patients had conjunctival congestion as the initial symptom. The average duration of conjunctival congestion was 5.9 ± 4.5 days (mean [SD]). The other ocular symptoms, including increased conjunctival secretion, ocular pain, photophobia, dry eye and tearing, were also found in patients with conjunctival congestion. Notably, hand-eye contact was independently correlated with conjunctival congestion in COVID-19 patients. We also found that some COVID-19 patients had chronic eye diseases, including conjunctivitis (33, 6.2%), xerophthalmia (24, 4.5%) and keratitis (14, 2.6%). Similar to the published studies, the most common clinical symptoms were fever, cough and fatigue. A total of 343 patients (64.1%) had positive SARS-CoV-2 detection in nasopharyngeal swabs.
- **Conclusions:** Conjunctival congestion is one of the COVID-19-related ocular symptoms, which could occur as the initial symptoms. Frequent hand-eye contact may be the risk factor for conjunctival congestion in COVID-19 patients. Screening of patients with conjunctival congestion by ophthalmologists is advocated during the outbreak of COVID-19. It is essential to provide eye-care equipment and strengthen education on eye protection.
Lien vers l'article: <https://pubmed.ncbi.nlm.nih.gov/32421258/>

Etude transversale par questionnaire

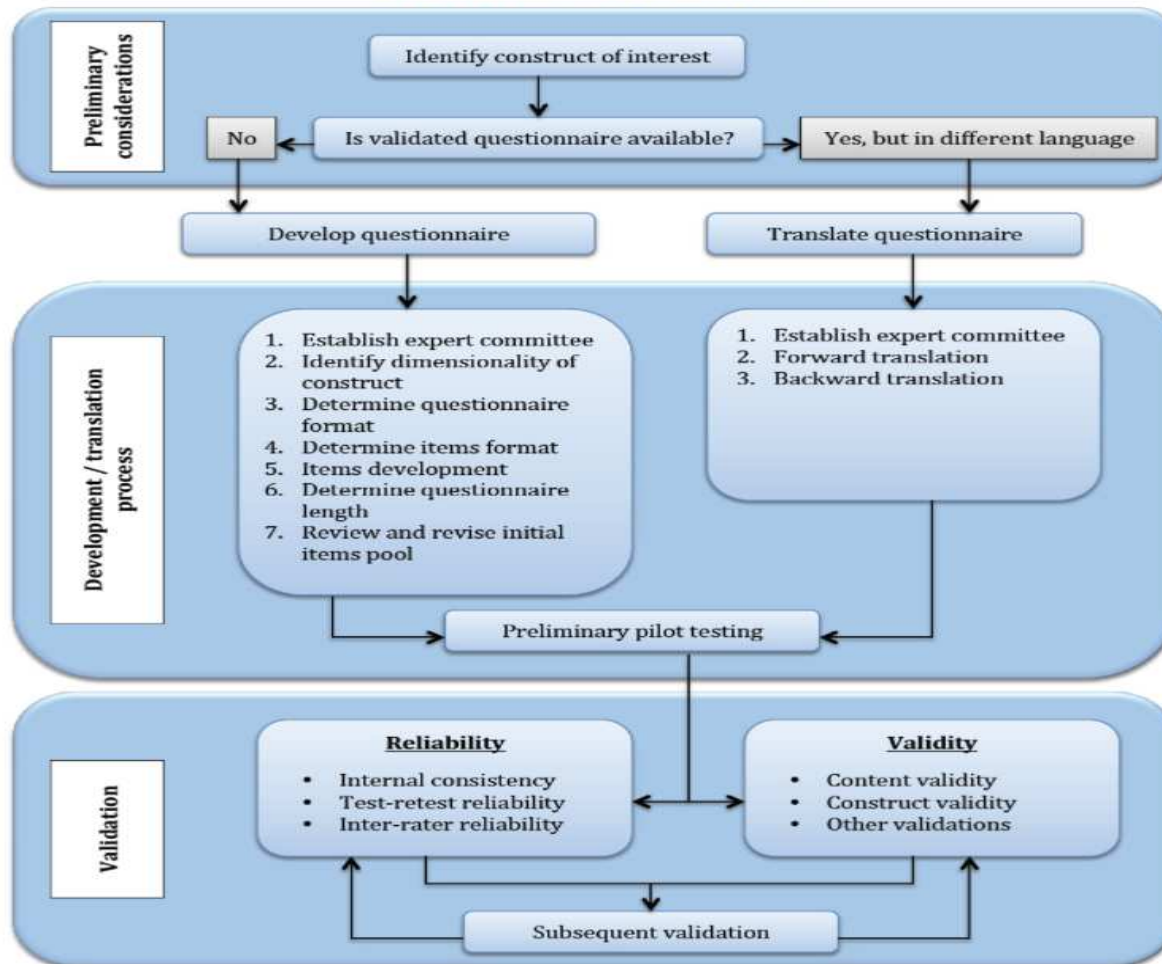
- utile pour comprendre comment les gens pensent;
- on peut mesurer:
 - attitudes,
 - des avis
 - comportements
- une série de questions liées au même construction / aspect d'intérêt;
- couvrant différents aspects;
- les réponses peuvent être combinées pour former un score ou une échelle;
- les échelles peuvent mesurer la gravité de la maladie ou elles peuvent être utilisées comme des tests de diagnostic

Etude transversale par questionnaire

- la qualité d'un questionnaire:
 - Analyse de fiabilité= précision de mesure et la reproductibilité des résultats
 - analyse de **cohérence interne**= pour déterminer dans quelle mesure les questions appartenant à la même échelle mesurent la même chose (les questions sont liés les uns aux autres)
 - analyse de **stabilité** (ex. **test-retest reliability**)= niveau de correspondance entre deux mesures distinctes effectuées par le même chercheur avec le même questionnaire/item
 - Analyse de validité = vérifier que le questionnaire mesure bien ce qu'il est censé évaluer
 - validité critérielle (*engl. Criterion validity*)
 - validité du construit (*engl. Construct validity*)
 - autres type de validité

Étapes de développement du questionnaire

Figure 1: Questionnaire development and translation processes



Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi anesthesia. 2017;11(5): 80-89.

Lien vers l'article:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5463570/>

Exemple

Lien vers l'article:

<https://pubmed.ncbi.nlm.nih.gov/32226353/>

International Journal of Mental Health and Addiction
<https://doi.org/10.1007/s11469-020-00270-8>

BRIEF REPORT

The Fear of COVID-19 Scale: Development and Initial Validation

Daniel Kwasi Ahorsu¹ · Chung-Ying Lin¹ · Vida Imani² · Mohsen Saffari³ · Mark D. Griffiths⁴ · Amir H. Pakpour^{5,6}

Published online: 27 March 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020



- **Abstract**
- **Background:** The emergence of the COVID-19 and its consequences has led to fears, worries, and anxiety among individuals worldwide. The present study developed the Fear of COVID-19 Scale (FCV-19S) to complement the clinical efforts in preventing the spread and treating of COVID-19 cases.
- **Methods:** The sample comprised 717 Iranian participants. The items of the FCV-19S were constructed based on extensive review of existing scales on fears, expert evaluations, and participant interviews. Several psychometric tests were conducted to ascertain its reliability and validity properties.
- **Results:** After panel review and corrected item-total correlation testing, seven items with acceptable corrected item-total correlation (0.47 to 0.56) were retained and further confirmed by significant and strong factor loadings (0.66 to 0.74). Also, other properties evaluated using both classical test theory and Rasch model were satisfactory on the seven-item scale. More specifically, reliability values such as internal consistency ($\alpha = .82$) and test-retest reliability (ICC = .72) were acceptable. Concurrent validity was supported by the Hospital Anxiety and Depression Scale (with depression, $r = 0.425$ and anxiety, $r = 0.511$) and the Perceived Vulnerability to Disease Scale (with perceived infectability, $r = 0.483$ and germ aversion, $r = 0.459$).
- **Conclusion:** The Fear of COVID-19 Scale, a seven-item scale, has robust psychometric properties. It is reliable and valid in assessing fear of COVID-19 among the general population and will also be useful in allaying COVID-19 fears among individuals.

Etudes écologiques corrélationnelles

- Les **études écologiques corrélationnelles** = étude dans laquelle les facteurs étudiés proviennent de population et non pas d'individu, définie à partir de caractéristiques temporelles ou géographiques.
- utiliser des données déjà collectées et agrégées au niveau de la (des) population (s).
- La mesure de l'association entre les variables observées peut être déterminé par le coefficient de corrélation de Pearson (r)

Etudes écologiques: **Avantages et désavantages**

- **Avantages:**

- Rapide, peu onéreux.
- De nouvelles hypothèses peuvent être générées
- De nouveaux facteurs de risque peuvent être identifiés

- **Désavantages:**

- Impossibilité de contrôler les facteurs de confusion
- Erreur écologique = erreur de raisonnement dans l'interprétation de résultats statistiques au niveau individuel à partir de données agrégées (conclusions fausses qui laissent penser que les relations entre des variables observées au niveau des groupes s'appliquent aussi à chaque individu)

Indicateurs calculés dans les études descriptives transversales

- **prévalence** = le rapport entre le nombre total de cas de la maladie à un moment donné (prévalence instantanée) ou sur une période de temps (prévalence périodique) et la taille de la population.
- **incidence** = le rapport entre le nombre de nouveaux cas de maladie sur une période donnée et la taille de la population.
- **taux de mortalité** = le rapport entre le nombre de décès enregistrés et la taille de la population sur une certaine période de temps.
- **taux de létalité** = le rapport entre le nombre de décès enregistrés et le nombre total de sujets malades sur une période donnée.

Exemple

> Nicotine Tob Res. 2020 Aug 24;22(9):1646-1649. doi: 10.1093/ntr/ntaa121.

Smoking Prevalence and COVID-19 in Europe

Panagiotis Tsigaris¹, Jaime A Teixeira da Silva²

FULL TEXT LINKS



Abstract

- **Introduction:** This ecological study investigates the [association between smoking prevalence and COVID-19 occurrence and mortality in 38 European nations](#) as of May 30, 2020.
- **Methods:** Data were collected from Our World in Data. Regression analysis was conducted to adjust for potential confounding factors such as economic activity (gross domestic product), the rate of COVID-19 testing, and the stringency of COVID-19 control policies.
- **Results:** There was a statistically significant negative association between smoking prevalence and the prevalence of COVID-19 across the 38 European nations after controlling for confounding factors ($p = 0.001$). A strong association was found between the prevalence of COVID-19 per million people and economic activity ($p = 0.002$) and the rate of COVID-19 testing ($p = 0.0006$). Nations with stricter policy enactment showed fewer COVID-19 cases per million people, but the association was not significant ($p = 0.122$). Delaying policy enactment was associated with a greater prevalence of COVID-19 ($p = 0.0535$). Evidence of a direct association between smoking prevalence and COVID-19 mortality was not found ($p = 0.626$). There was a strong positive association between COVID-19 mortality rate and the prevalence of COVID-19 cases ($p < 0.0001$) as well as the proportion of the population over 65 years of age ($p = 0.0034$) and a negative association with the rate of COVID-19 testing ($p = 0.0023$).
- **Conclusions:** We found a negative association between smoking prevalence and COVID-19 occurrence at the population level in 38 European countries. This association may not imply a true or causal relationship, and smoking is not advocated as a prevention or treatment of COVID-19.

Lien vers l'article: <https://pubmed.ncbi.nlm.nih.gov/32609839/>

Modélisation des phénomènes de santé à l'aide de la régression

Régression

- **But:** - méthode pour étudier liaison entre 2 ou plusieurs variables

- **Forme générale:**

$Y = f(X) + \varepsilon$ ou f = fonction de régression

X = variable indépendante, explicative, prédictive

Y = variable dépendante, expliquée, prédite

ε = écart (erreur) de cette approximation

Type de Régression

- le type du variable dépendante
 - variable quantitative – régression linéaire
 - variable dichotomique – régression logistique
- la linéarité de la fonction
 - régression linéaire
 - régression non linéaire
- le nombre de variables dépendantes:
 - régression univariée (une variable dépendante)
 - régression multivariée (≥ 2 variables dépendantes)
- le nombre de variables indépendantes:
 - régression **simple** (une variable indépendante)
 - régression **multiple** (≥ 2 variables indépendantes)

Régression et Causalité

- Le fait qu'on trouve dans une étude une dépendance statistiquement significative, **ne signifie pas fortement que la relation est causale !!!**
- Ces informations, nous indiquent qu'il y a des relations, entre les variables, mais on ne sait pas qui est la cause et qui est l'effet.
- Pour montrer la causalité, la méthodologie de la réalisation de l'étude est très importante.

Regression logistique

- **But:** évaluer la liaison entre **1 variable qualitative** (dichotomique/nominale/ordinaire) **dépendante** et ≥ 1 **variables** quantitatives ou qualitatives **indépendantes**
 - Deux **objectifs** possibles:
 - on cherche d'expliquer la survenue (mesurer par l' odds /probabilité) d'un **événement binaire (oui/non) dépendante**: tester d' hypothèses
 - on cherche de prédire la survenue (mesurer par l' odds /probabilité) d'un événement binaire (oui/non) dépendante: model prédictif
- en présence de covariables = toute variable indépendante associée à la variable dépendante dans un modèle de régression.

Regression logistique

- Très populaire en médecine: la variable dépendante peut être l'apparition d'une maladie, et les variables dépendantes sont des facteurs de risque.
- La variable dépendante doit être une **variable qualitative dichotomique**
 - Présence de maladie ou de décès,
 - Diabète (oui / non)
 - Hypertension (oui / non)
- **Variables indépendantes** : ≥ 1 variables de différents types
 - **Qualitatives**: Consommation d'alcool (oui/non), tabagisme (oui/non), obésité (oui/non)
 - **Qualitatives**: Consommation d'alcool (oui/non), tabagisme (oui/non), obésité (oui/non)
- *Conditions d'applicabilité*
- les observations sont indépendantes les unes des autres
- La relation entre chaque prédicteur continue et la variable dépendante est linéaire sur l'échelle logit

Model de regression logistique **simple**

X_1 –variable indépendante et Y =variable dépendante

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = b_0 + b_1 X_1$$

ou Odds (ou chance relative)

$p = \Pr(Y=1/X_1)$ la probabilité de survenu d'un événement en fonction de la valeur prise par la variable indépendante (X_1)

b_0 = le terme libre (constante du modèle)

b_1 = coefficient de régression partielle -> le rapport des chances (OR)

Model de regression logistique simple

- Odds ($Y=1/X_1 = 0$ =absente) = e^{b_0} ou e = constante mathématique, également appelée nombre d'Euler
- Odds ($Y=1/X_1 = 1$ =present) = $e^{b_0+b_1}$
- l'odds ratio pour $Y=1 \rightarrow OR = \frac{\text{Odds}(Y=1/X_1 = 1)}{\text{Odds}(Y=1, X_1 = 0)} = \frac{e^{b_0+b_1}}{e^{b_0}} = e^{b_1}$

Model de regression logistique **multiple**

X_1, X_2, \dots, X_k -variables independantes et Y =variable dépendante

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$$

ou

$p = \Pr(Y=1/X_1, X_2, \dots, X_k)$ = la probabilité de survenu d'un enveniment en fonction des valeurs prise par les variables indépendantes (X_1, X_2, \dots, X_k)

b_0 = le terme libre (constante du modèle)

b_1, \dots, b_k = coefficient de régression logistique -> le rapport des chances (OR)

Model de regression logistique multiple

Interprétation de l'odds ratio (OR) du modèle logistique quand la variable indépendante (X_i) est:

- **dichotomique:**

$$OR = \frac{\textit{l'odds de l'event pour group exposé}}{\textit{l'odds de l'event pour group non exposé (de reference)}}$$

si $b_i > 0$ alors $OR > 1$ (facteur de risk),

si $b_i < 0$ alors $OR < 1$ (facteur protecteur).

- **nominal avec p categories:**

- On obtient « p-1 » des OR (la variable nominale doit être transformée en variable dichotomique p-1 avant la régression)

- **quantitative continues:**

- représente le changement de l'odds de l'événement d'intérêt lorsque X_i augmente d'une unité après ajustement pour les autres prédicteurs

La performance du modele logistique

- Le test de Hosmer – Lemeshow ne devrait pas être significatif
 - Hypothese nulle H_0 : Le modèle testé est adéquat pour les données.
- précision de la prédiction (en %)
- Calcul des coefficients de McFadden R^2 , Cox-Snell R^2 , Nagelkerke / Cragg & Uhler's R^2

Quantification de l'importance de la relation pour plusieurs facteurs

Régression logistique **simple**, **multiple**, odds ratio crue/ajusté...

- Une régression logistique avec une seule variable indépendante (**régression logistique simple**)
 - permet une analyse univariée – on parle d'un régression logistique simple
 - nous offre un odds ratio brut/crue (« crude »/ » unadjusted» en anglais - sans ajustement, pour la variable indépendante
- Une régression logistique avec une **plusieurs variables indépendantes** (**régression logistique multiple**)
 - permet une analyse multivariée – on parle d'un régression logistique multiple
 - nous offre un odds ratio ajusté (« adjusted » en anglais) pour la variable indépendante – avec ajustement pour les autres variables indépendants dans la régression
- L'odds ratio **ajusté** peut nous rapprocher **plus** a la vérité que le odds ratio brut (« crude »/ » unadjusted» – obtenu a l'aide d'une régression logistique simple, ou calcul $a*d/b*c$), **parce que on peut tenir compte (on peut contrôler) d'autres facteurs de risque qui agit dans le même temps**

Quantification de l'importance de la relation pour plusieurs facteurs

Interprétation statistique du odds ratio crue/brut/non ajuste d'une variable qualitative dichotomique pour un régression logistique simple (uni variée)

- L'équation du **régression logistique simple (uni variée)**, avec un **variable indépendante qualitative dichotomique**:
- $\text{Ln (odds (Maladie =oui=1))} = b_0 + b_1 * X_1 \text{ (oui=1/non=0)} \rightarrow$
- $\text{L'odds (Maladie =oui=1)} = e^{b_0} * e^{b_1 * X_1} \text{ (oui=1/non=0)} \rightarrow \text{OR}_1 = \frac{\text{Odds (Y=1/X}_1 = 1)}{\text{Odds (Y=1,X}_1 = 0)}$
- **L'odds ratio** (Maladie =oui) = OR_1 associé au facteur X_1 (oui/non)
- Ex: **L'odds ratio** (cheilopalatoschizis =oui) = $\text{OR}_1 = 1,90$ associé au facteur Alcool (oui/non)
- L'interprétation du **Odds ratio crue/brut, sans ajustement**
- Si la variable est **qualitatives dichotomiques** (ex. consommation d'alcool):
 - il y a 1,90 fois plus de la chance (l'odds) d'avoir la maladie - cheilopalatoschisis pour ceux avec exposition – avec consommation d'alcool comparée a ceux qui n'ont pas été exposée a l'alcool.

Quantification de l'importance de la relation pour plusieurs facteurs

Interprétation statistique du odds ratio crue/brut/non ajusté d'une variable quantitative pour un régression logistique simple (uni variée)

- L'équation du régression logistique simple (uni variée), avec un **variable indépendante quantitative**:
- $\text{Ln (odds (Maladie =oui))} = b_0 + b_2 * X_2$ (quantitative)
- $\text{L'odds (Maladie =oui)} = e^{b_0} * e^{b_2 * X_2} \rightarrow \text{OR}_2 = \frac{\text{Odds (Y=1/X}_2 = a+1)}{\text{Odds (Y=1, X}_2 = a)}$ ou a=valeur fixé de la variable X_2
- **L'odds ratio** (Maladie =oui) = OR_2 associé au facteur X_2
- Ex: **L'odds ratio** (cheilopalatoschizis* =oui) = $\text{OR}_2 = 1,10$ associé au facteur Age de la mère (années)
- L'interprétation **de l' Odds ratio crue/brut, sans ajustement**
- Si la variable est quantitative (ex. âge de la mère):
 - pour chaque unité de mesure de la variable (pour chaque **année de la mère**) en plus la chance (l'odds) d'avoir cheilopalatoschisis augmente **1,10 fois**

* syndrome des "lèvres du lapin"

Régression logistique simple- exemple interprétation

Caractéristique	OR crude (95% CI)	P-value
Alcool pendant la grossesse (oui/non)	2,25 (1,64 – 7,83)	<0,001
Age de la mère (années)	1,22 (1,03 – 1,85)	0,03

Régression logistique uni variée pour la prédiction de l'odds du cheilopalatoschizis.

L'OR – odds ratio - **crude – brut - unadjusted** (voir tableau) – (brut signifie qu'il s'agit - entre deux variables seulement – une variable dépendent (le cheilopalatoschizis) et **une variable indépendante** (la consommation d'alcool pendant la grossesse) - d'un analyse **uni variée**. Ici dans le tableau il y a deux régressions uni variées, entre cheilopalatoschizis et la consommation d'alcool pendant la grossesse , et une entre cheilopalatoschizis et l'âge de la mère

Pour la variable dichotomique: - OR crude – brut - unadjusted pour la consommation d'alcool pendant la grossesse : il y a 2,25 fois plus de chance de cheilopalatoschizis pour ceux qui ont été expose a l'alcool par rapport au ceux qui n'ont pas été exposée.

Pour la variable quantitative: - OR crude – brut - unadjusted pour l'âge de la mère: il y a 1,22 fois plus de chance de cheilopalatoschizis pour chaque unité de mesure en plus de l'âge de la mère (pour chaque année)

Régression logistique simple & multiple– exemple interprétation

Caractéristique	OR crude (95% CI)	P-value	OR adjusted (95% CI)	P-value
Alcool pendant la grossesse (oui/non)	2,25 (1,64 – 7,83)	<0,001	2,10 (1,52– 9,95)	0,001
Age de la mère (années)	1,22 (1,03 – 1,85)	0,03	1,15 (1,02 – 1,78)	0,046

Régression logistique **multiple (multi variée)**, pour la prédiction de l'odds du cheilopalatoschizis.

OR – odds ratio - ajusté - adjusted (voir tableau) – (ajusté signifie qu'il s'agit d'une analyse entre une variable dépendante (le cheilopalatoschizis) et **plusieurs variables indépendantes** (consommation d'alcool pendant la grossesse et l'âge de la mère)– analyse **multivariée**

Variable dichotomique: - OR ajusté - adjusted pour la consommation d'alcool pendant la grossesse : il y a 2,10 fois plus de chance de cheilopalatoschizis pour ceux qui ont été exposés à l'alcool par rapport à ceux qui n'ont pas été exposés, on tenant constantes (en tenant fixe/en ajustant/ en contrôlant) les autres variables (âge de la mère).

Variable quantitative: - OR ajusté - adjusted pour l'âge de la mère: il y a 1,15 fois plus de chance de cheilopalatoschizis pour chaque unité de mesure en plus de l'âge de la mère (pour chaque année) , on tenant constantes (en tenant fixe/en ajustant/ en contrôlant) les autres variables (consommation d'alcool pendant la grossesse)

Analyse multivariée: la Quantification de l'importance de la relation pour plusieurs facteurs

Interprétation statistique du odds ratio ajusté d'une variable qualitative dichotomique, et d'une variable quantitative pour une régression logistique multiple (multivariée)

- L'équation du régression logistique multiple (multivariée):
- $\ln(\text{odds}(\text{Maladie}=\text{oui})) = b_0 + b_1 * X_1 + b_2 * X_2$
- Nous modélisons l'odds de la maladie à l'aide de deux odds ratios ajustés (OR₁ ajusté et OR₂ ajusté)
- L'interprétation du **Odds ratio ajusté** pour des variables
 - **Qualitatives dichotomiques** (ex. consommation d'alcool): $\text{OR}_1 = \text{ajusté} \frac{\text{Odds}(Y=1/X_1=1, X_2)}{\text{Odds}(Y=1, X_1=0, X_2)}$ → il y a 2,10 fois plus de la chance d'avoir la maladie - cheilopalatoschisis pour ceux avec exposition – avec consommation d'alcool comparée à ceux qui n'ont pas été exposés à l'alcool si on tient l'autre variable constante (on ajuste pour la 2ème variable ici l'âge)
 - **Quantitatives** (ex. âge de la mère): $\text{OR}_2 = \text{ajusté} \frac{\text{Odds}(Y=1/X_2=a+1, X_1)}{\text{Odds}(Y=1, X_2=a, X_1)}$ pour chaque unité de mesure de la variable (pour chaque année de la mère) en plus la chance d'avoir cheilopalatoschisis augmente 1,15 fois si on tient l'autre variable constante/fixé (on ajuste pour l'autre variable - ici l'alcool)

Exemple

> Can J Infect Dis Med Microbiol. 2021 May 7;2021:5515941. doi: 10.1155/2021/5515941. eCollection 2021.

Risk Factors for COVID-19 in Patients with Hypertension

Xinxin Wang ¹, Haihua Zhang ¹, Huan Du ¹, Ruina Ma ^{1 2}, Yandong Nan ^{1 2}, Tao Zhang ^{1 2}

Abstract

- **Background:** Hypertension, as the most common comorbidity for patients with coronavirus disease 19 (COVID-19), has resulted in cases with more severe symptoms and higher mortality. The risk factors associated with COVID-19 in patients with hypertension are unknown.
- **Methods:** All the available and confirmed patients with COVID-19 from February 3 to March 10, 2020, were enrolled from Huoshenshan Hospital, Wuhan, China. The demographic characteristics, clinical manifestations, laboratory data, radiological assessments, and treatments on admission were extracted and compared. Univariate and multivariate logistic regression methods were used to explore risk factors associated with COVID-19 in patients with hypertension and the severity of the cohort.

Lien vers l'article: <https://pubmed.ncbi.nlm.nih.gov/34055114/>

Exemple: Modelés de régression logistique simple et multiple

- Résultats partiels extraits de l'article;

TABLE 4: Univariate logistic regression analysis of patients with COVID-19 and hypertension.

	Univariate OR (95% CI)	p value
Diabetes	2.148 (1.074–4.297)	0.031
Coronary heart disease	3.009 (1.344–6.733)	0.007
Cerebrovascular disease	4.279 (1.181–15.510)	0.027
Percentage of PIV, %	1.025 (1.004–1.046)	0.021
White blood cell count, $\times 10^9/L$	1.133 (1.035–1.241)	0.007
Neutrophil count, $\times 10^9/L$	1.128 (1.032–1.232)	0.008
ALP, IU/L	1.007 (0.998–1.016)	0.112
Urea nitrogen, mmol/L	1.142 (1.037–1.257)	0.007
UA (uric acid), $\mu\text{mol/L}$	1.003 (1–1.005)	0.028
Creatinine, $\mu\text{mol/L}$	1.016 (1.004–1.027)	0.008
K ⁺	0.582 (0.383–0.885)	0.011

ALT: alanine aminotransferase; OR: odds ratio; CI: confidence interval.

Analyse univarié de la régression logistique (la régression logistique simple)

L'estimation ponctuelle de l'odds ratio non-ajusté (brut/crue) & 95% intervalle de confiance (IC)

TABLE 5: Multivariate logistic regression analysis of patients with COVID-19 and hypertension.

	Multivariate OR (95% CI)	p value
Diabetes	2.438 (0.824–7.210)	0.107
Coronary heart disease	5.281 (1.462–19.076)	0.011
Cerebrovascular disease	5.661 (0.952–33.662)	0.057
Percentage of PLV, %	0.994 (0.963–1.026)	0.71
Neutrophil count, $\times 10^9/L$	1.471 (1.183–1.828)	0.001
Urea nitrogen, mmol/L	0.994 (0.795–1.244)	0.961
UA (uric acid), $\mu\text{mol/L}$	1.003 (0.998–1.008)	0.247
Creatinine, $\mu\text{mol/L}$	1.018 (0.994–1.004)	0.149
K ⁺	0.273 (0.134–0.555)	<0.001

OR, odds ratio; CI, confidence interval.

Analyse multivarié de la régression logistique (le model multivariable de la régression logistique multiple)

L'estimation ponctuelle de l'odds ratio ajusté & 95% intervalle de confiance (IC)

MERCI POUR VOTRE ATTENTION!!