**Ma partie** : **brouillon de l’article de statistiques**

**Bibliographie** :

<https://www.sante.fr/anevrisme-0>

<https://www.chuv.ch/fr/neurochirurgie/nch-home/patients-et-familles/maladies-traitees/lanevrisme-cerebral>

<https://www.vidal.fr/maladies/coeur-circulation-veines/anevrisme-anevrysme.html>

<https://www.elsan.care/fr/pathologie-et-traitement/maladies-vasculaires/anevrisme-definition>

**Titre** : Recherche de différents groupes de patients présentant un risque de récidive d’anévrisme en fonction de plusieurs paramètres cliniques et démographiques, tels que l’indice de masse corporelle, l’âge, le sexe et le pays d’origine, et prédiction du niveau de risque de récidive de la maladie selon ces précédents paramètres.

**Glossaire** :

*IMC* : Indice de masse corporelle

**Introduction** :

Les anévrismes s’apparentent à une dilatation localisée d’artères, entraînant au long terme une fragilisation du vaisseau, avec une éventuelle création de caillots sanguins, ainsi qu’une possible rupture. Ce phénomène, appelé rupture d’anévrisme, cause alors une hémorragie interne. Toute artère, de taille variable dans l’organisme, peut être affectée. Les plus fréquents touchent les artères du cerveau, cause importante d’accident vasculaire cérébral (AVC), et l’aorte.

En prenant le cas cérébral, 2 à 4 % de la population mondiale serait porteuse d’un anévrisme. Le risque de récidive est alors important quelques jours après l’hémorragie, et atteint les 40 à 50 % les six mois suivants. Ce problème majeur de santé publique est ainsi susceptible de se produire chez plusieurs types de personnes, catégorisées selon le Vidal. Parmi les personnes les plus à risque d’anévrisme et par conséquent de récidive, se retrouvent entre autres les personnes atteintes d’obésité et les personnes d’âge avancé. D’autre part et selon la même source, les hommes sont caractérisés comme treize fois plus susceptibles de développer un anévrisme de l’aorte, le plus généralement après 65 ans. Toutefois, en prenant le cas de l’anévrisme cérébral et selon le site du service de neurochirurgie du CHUV de Lausanne, les femmes développent un peu plus souvent des anévrismes cérébraux.

Ainsi, au regard d’une étude statistique réalisée en France et au Vietnam, le risque de récidive d’anévrisme chez des personnes, déjà prises en charge de la maladie, a été évalué sur une population de 209 individus. Il serait alors pertinent d’analyser, au regard des conclusions tirées des articles scientifiques précédemment cités, si des paramètres comme l’indice de masse corporelle, l’âge et le sexe de la personne peut effectivement expliquer le niveau de risque de récidive d’anévrisme, et si nous pouvons prédire le niveau de risque en fonction d’une de ces données cliniques. Nous pouvons alors étudier les deux questions suivantes :

-          ***S’il nous pouvons identifier plusieurs groupes de patients présentant un risque de récidive selon les paramètres comme l’indice de masse corporelle, l’âge, le sexe et l’origine démographique, existe-t-il également un lien avéré entre ces paramètres explicatifs et le risque de récidive d’anévrisme, ?***

-        ***Si ces groupes trouvés sont associés à des niveaux distincts de risque de récidive d’anévrisme, est-il possible de prédire le niveau de risque de récidive selon ces données cliniques et démographiques du patient ?***

L’indice de masse corporelle est également caractérisé comme le carré de la taille d’un individu, en mètre, divisé par son poids, en kilogramme.

**1️⃣ Analyse de groupes de clusters selon le BMI**

**📌 Objectif : Identifier des groupes de patients selon leur indice de masse corporelle (BMI).**

Tu peux utiliser **Mclust** (classification gaussienne) pour regrouper les patients en fonction de leur BMI :

r

CopyEdit

library(mclust)

# Classification gaussienne sur le BMI

d.mc\_bmi <- Mclust(d$bmi, G=3:7) # On cherche entre 3 et 7 groupes optimaux

summary(d.mc\_bmi) # Vérifier combien de clusters sont retenus

plot(d$bmi, col=d.mc\_bmi$classification, pch=20, main="Clusters sur le BMI")

🔹 **Sortie attendue :** Un nombre optimal de clusters (d.mc\_bmi$G) et une visualisation des groupes formés.

**2️⃣ Voir si le risque de récidive est influencé par le BMI**

**📌 Objectif : Vérifier si les groupes BMI sont associés au nombre de récidives.**

On peut tester cela avec une **ANOVA** (car risk est une variable numérique) :

r

CopyEdit

# ANOVA pour voir si le risque diffère entre les groupes BMI

anova\_bmi <- aov(d$risk ~ d.mc\_bmi$classification, data=d)

summary(anova\_bmi)

🔹 **Interprétation :**

* Si la **p-value** est **< 0.05**, alors les groupes BMI influencent significativement le risque de récidive.
* On peut visualiser l'effet avec un **boxplot** :

r

CopyEdit

boxplot(d$risk ~ d.mc\_bmi$classification, main="Nombre de récidives par groupe BMI", xlab="Groupe BMI", ylab="Risque de récidive")

3)

**Option 2 : Modèle de régression linéaire (si risk est une variable continue)**

Si risk est un **nombre de récidives (0, 1, 2, …)** et pas une variable binaire, on peut utiliser une **régression linéaire** :

r

CopyEdit

model\_lm <- lm(risk ~ bmi, data=d)

summary(model\_lm)

* On regarde si **bmi est significatif** (p-value < 0.05).
* On peut **prédire le risque moyen** pour un patient avec un certain BMI :

r

CopyEdit

d$pred\_risk <- predict(model\_lm)

plot(d$bmi, d$risk, pch=20, col="blue", main="Risque prédit en fonction du BMI")

lines(d$bmi, d$pred\_risk, col="red", lwd=2) # Ajout de la courbe de tendance

**Option 3 : Classification avec une matrice de confusion**

Si tu veux **catégoriser** le risque en faible/moyen/élevé, tu peux définir des seuils et utiliser une **matrice de confusion** :

r

CopyEdit

d$risk\_cat <- cut(d$risk, breaks=c(-1, 0, 2, Inf), labels=c("Faible", "Moyen", "Élevé"))

# Modèle de classification (exemple : arbre de décision)

library(rpart)

model\_tree <- rpart(risk\_cat ~ bmi, data=d)

plot(model\_tree)

text(model\_tree, use.n=TRUE)

# Prédiction et matrice de confusion

d$pred\_risk\_cat <- predict(model\_tree, type="class")

table(d$risk\_cat, d$pred\_risk\_cat) # Matrice de confusion