

# Essai clinique de médicament sur la cholangite biliaire primitive

## Contexte des données

La cholangite biliaire primitive (CBP), nouvelle dénomination de la cirrhose biliaire primitive, est une maladie cholestatique auto-immune et représente la première cause de cholestase (ralentissement de la circulation biliaire) intra-hépatique. Elle se caractérise par une destruction des canaux biliaires de petite taille liée à une infiltration lymphocytaire. La prévalence de la CBP est de 10 à 40 pour 100 000 habitants. La CBP affecte préférentiellement la femme (dans 90 % des cas) d'un âge médian de 55 ans au diagnostic. Ainsi, chez la femme de plus de 40 ans, la prévalence de la CBP est de 1 pour 1 000. Les signes biologiques de la maladie sont les suivants : il existe une anomalie portant sur les phosphatases alcalines, les gamma-GT, ou les transaminases. Ces enzymes sont des protéines présentes dans le foie. Lorsque la sécrétion biliaire est anormale du fait de l'inflammation des voies biliaires, ces enzymes sont fabriquées en excès par le foie et déversées dans le sang, si bien que leur concentration sanguine augmente au-dessus des valeurs normales.

Le jeu de données fourni ici suit un essai clinique randomisé ayant eu lieu entre 1974 et 1986 pour tester l'effet de l'utilisation d'un médicament (la D-penicillamine) sur la CBP. Il porte sur 312 patients pour lesquels la majorité de biomarqueurs potentiels est connue. Des données indépendantes de 106 patients supplémentaires ont été ajoutées afin de valider le modèle.

## Descriptif des données

20 colonnes et 418 lignes :

Nom de la colonne	Type de variable	Description
ID	Nombre entier	Numéro d'identification du patient
N_Days	Nombre entier	Nombre de jours entre l'entrée dans l'étude et soit le décès, la greffe de foie ou la fin de l'étude en juillet 1986
Status	Chaîne de caractères	Statut du patient : C = sortie de l'étude (= survie), CL = sortie de l'étude du fait d'une greffe de foie, D = décès
Drug	Chaîne de caractères	Traitement administré durant l'essai clinique : D-penicillamine = D-penicillamine, Placebo = placebo
Age	Nombre entier	Âge en jours
Sex	Chaîne de caractères	Sexe biologique : F = femme, M = homme
Ascites	Chaîne de caractères	Présence d'ascites, épanchement de liquide dans l'abdomen : N = non, Y = oui
Hepatomegaly	Chaîne de caractères	Présence d'hépatomégalie, c'est-à-dire d'hypertrophie du foie : N = non, Y = oui
Spiders	Chaîne de caractères	Présence d'angiome stellaire ( <i>spiders</i> en anglais) : N = non, Y = oui
Edema	Chaîne de caractères	Présence d'oedème : N = non, Y = oui
Bilirubin	Nombre réel	Bilirubine sérique en mg · dL <sup>-1</sup>
Cholesterol	Nombre entier	Cholestérol sérique en mg · dL <sup>-1</sup>
Albumin	Nombre réel	Albumine en mg · dL <sup>-1</sup>

Nom de la colonne	Type de variable	Description
Copper	Nombre entier	Cuivre dans l'urine en mg · dL <sup>-1</sup>
Alk_Phos	Nombre réel	Phosphatase alcaline en unité arbitraire · L <sup>-1</sup>
SGOT	Nombre réel	Aspartate aminotransférase sérique en unité arbitraire · mL <sup>-1</sup> (en anglais, <i>serum glutamic oxaloacetic transaminase</i> )
Triglycerides	Nombre entier	Triglycérides en mg · dL <sup>-1</sup>
Platelets	Nombre entier	Nombre de plaquettes sanguine en milliers · ml <sup>-3</sup>
Prothrombin	Nombre réel	Taux de prothrombine (= indicateur de la coagulation sanguine) exprimé en pourcentage du temps de Quick
Stage	Nombre entier	Stade histologique de la maladie : 1 = stade I, 2 = stade II, 3 = stade III, 4 = stade IV (voir explications ci-dessous)

## Explications de la variable Stage

Les stades I, II, III et IV correspondent respectivement aux stade portal, périportal, septal (ou lobulaire), et à la cirrhose. Les stades III et IV correspondent aux stades avancés de la maladie.

## Travail demandé

### Rapport à rendre

Vous réaliserez un court document (4 à 6 pages), introduisant le jeu de données et les problématiques que vous vous êtes posées. Vous ajouterez l'ensemble de votre notebook d'analyse en annexe de votre rapport.

Il n'est pas nécessaire d'analyser toutes les données présentes, vous pouvez vous concentrer sur un sous-ensemble de données.

Votre rapport devra être structuré de la manière suivante :

1. Introduction (présentant le contexte de votre étude et votre problématique)
2. Préparation des données (présentant des analyses descriptives des données et vos éventuels filtres sur les données)
3. Résultats (vous réalisez vos analyses et présentez vos résultats)
4. Discussion (vous recontextualisez vos résultats par rapport à votre problématique, vous en présentez les limites principales) avec une partie “Implications pour la pratique officinale” (que signifient vos résultats pour la Pharmacie et votre future pratique ?)

## Référence

Article original : > Dickson, Rolland E. M.D.; Grambsch, Patricia M; Fleming, Thomas R; Fisher, Lloyd D.; Langworthy, Alice1. Prognosis in primary biliary cirrhosis: Model for decision making. Hepatology 10(1):p 1-7, July 1989. | DOI: <https://doi.org/10.1002/hep.1840100102>

Jeu de données : > Dickson, E., Grambsch, P., Fleming, T., Fisher, L., & Langworthy, A. (1989). Cirrhosis Patient Survival Prediction [Dataset]. UCI Machine Learning Repository. <https://doi.org/10.24432/C5R02G>.

Information générale sur la CBP : <https://www.filfoie.com/glossary/cholangite-biliaire-primitive-cirrhose-biliaire-primitive/>

Fiche synthétique de la Haute Autorité de Santé sur la CBP : [https://www.has-sante.fr/upload/docs/application/pdf/2021-11/synthese\\_mg\\_cholangite\\_biliaire\\_primitive.pdf](https://www.has-sante.fr/upload/docs/application/pdf/2021-11/synthese_mg_cholangite_biliaire_primitive.pdf)

Site d'information sur la CBP : <https://www.fmchgastro.org/texte-postu/postu-2018-paris/cholangite-biliaire-primitive/>