



IUT de Paris - Rives de Seine
Université Paris Cité

SAE Projet

Etude de l'impact de la cathétérisations cardiaque droite (RHC)

Matteo CAI
Mathias DE RIDDER
William LEFEBVRE
Makam TRAORE

Table des matières

1. Introduction	3
2. Présentation de données	4
2.1 Présentation des catégories.....	4
2.2 Variables utilisées.....	5
3. Analyse avancée : présentation une par une des problématiques et attentes déclinées dans l'introduction	11
3.1 Impact des variables sur DEATH	11
3.2 Impact de la technique RHC sur des variables spécifiques.....	17
4. Conclusion	23
5. Annexes	24

1. Introduction

La cathétérisations cardiaque droite, également connue sous le sigle RHC, est une procédure médicale cruciale utilisée pour évaluer et surveiller la fonction cardiaque chez les patients présentant une variété de pathologies cardiovasculaires. En insérant un cathéter dans les vaisseaux sanguins, notamment dans les cavités cardiaques droites, cette technique permet une évaluation précise des pressions, des débits sanguins et d'autres paramètres hémodynamiques vitaux. Depuis son introduction dans les années 1970, le RHC est devenu un outil essentiel en médecine, offrant des informations précieuses pour guider le diagnostic et la gestion des maladies cardiovasculaires complexes. Cependant, son utilisation nécessite une expertise technique et comporte des risques, ce qui soulève des questions concernant son application appropriée et son efficacité clinique.

Quels sont les principaux facteurs de risque de décès ? En quoi la cathétérisation cardiaque droite/ Right Heart Catheterization (RHC) influence-t-elle ces risques ?

Pour cette analyse, nous possédons une base de données provenant d'une étude observationnelle multicentrique réalisée aux Etats-Unis. Cette étude porte sur 5 500 patients adultes admis en unité de soins intensifs entre juin 1989 et janvier 1994, et suivis pendant 6 mois.

Dans un 1^{er} temps nous allons présenter les données et faire une analyse descriptive puis dans un 2nd temps analyser les variables facteurs de risque influençant le décès et pour finir voir l'impact spécifique de la cathétérisations cardiaque droite (RHC).

La base de données est composée de 67 variables, regroupées en 7 catégories distinctes. Ces dernières sont "Diagnostic à l'admission" reflétant les diverses affections médicales des patients, " Evénements" qui correspond au traitement et à l'état vital du patient, " Scores Cliniques", "Caractéristiques physiologiques", "Caractéristiques sociodémographiques", "Maladies associées" et "Dates".

2. Présentation de données

L'étude porte sur 5 500 patients adultes admis en unité de soins intensifs aux États-Unis entre juin 1989 et janvier 1994, suivis pendant 6 mois. Les données collectées sont regroupées en 67 variables réparties en 7 catégories distinctes : **Diagnostic à l'admission, Maladies associées, Scores Cliniques, Caractéristiques physiologiques, Événements, Dates, et Caractéristiques sociodémographiques.**

2.1 Présentation des catégories

Diagnostic à l'Admission

Cette catégorie inclut les principales affections médicales diagnostiquées lors de l'admission. Elle se divise en neuf sous-catégories. La plus fréquente est l'insuffisance rénale aiguë (ARF), représentant 36% des cas. Les problèmes respiratoires (RESP) et cardiovasculaires (CARD) suivent de près, chacun représentant 32% des cas. D'autres affections incluent le cancer (CA), les problèmes neurologiques (NEURO), gastro-intestinaux (GASTR), hématologiques (HEMA), infectieux (SEPS), et traumatiques (TRAUMA), bien que ces dernières soient moins fréquentes.

Maladies Associées

Les maladies associées sont diverses et souvent graves, influençant le pronostic des patients. Par exemple, l'immunosuppression (IMMUNHX), présente chez 17% des patients, inclut des cas de VIH et de diabète. D'autres conditions comme les hémorragies gastro-intestinales (GIBLEDHX) et les infarctus du myocarde (AMIHX) sont également notées, bien que leur prévalence soit plus faible, entre 3% et 11%.

Scores Cliniques

Les scores cliniques fournissent une évaluation standardisée de la gravité de l'état des patients. Le score DASI a une moyenne de 20.83, reflétant une variabilité considérable dans la capacité fonctionnelle des patients. Le score Apache 3, avec une médiane de 54.00 et des valeurs extrêmes atteignant 499.00, indique des niveaux de gravité très divers. Le score Glasgow présente une médiane de 0.00, révélant une distribution asymétrique avec de nombreux patients ayant un score de coma profond.

Caractéristiques Physiologiques

Les caractéristiques physiologiques des patients offrent des mesures précises de leur état physique. Le poids moyen des patients est de 74.49 kg, avec une plage allant de 19.50 kg à 244 kg. La température corporelle moyenne est de 37.62°C, bien que des extrêmes entre 27.00°C et 43.00°C soient observés. La tension artérielle moyenne varie de 0 mm Hg à 259 mm Hg, et la fréquence respiratoire moyenne est de 28.15 respirations par minute, avec des valeurs

extrêmes de 0 à 100. La fréquence cardiaque moyenne est de 115.3 battements par minute, avec une plage de 0 à 250 bpm, ce qui reflète des conditions de santé très diverses parmi les patients.

Événements

Les événements enregistrés incluent des interventions spécifiques et des résultats vitaux. Par exemple, 38.2% des patients ont bénéficié de la cathétérisations cardiaque droite (RHC). Le taux de mortalité parmi les patients étudiés est notablement élevé, atteignant 65.1%, ce qui souligne la gravité des conditions traitées en soins intensifs.

Dates

Les variables de dates sont essentielles pour suivre les événements clés durant la période de l'étude. Elles incluent la date d'inclusion dans l'étude, la date de sortie de l'hôpital, la date de décès éventuel, et la date du dernier contact avec le patient. Ces données temporelles permettent de suivre le parcours de chaque patient et d'analyser la durée des interventions et des séjours hospitaliers.

Caractéristiques Sociodémographiques

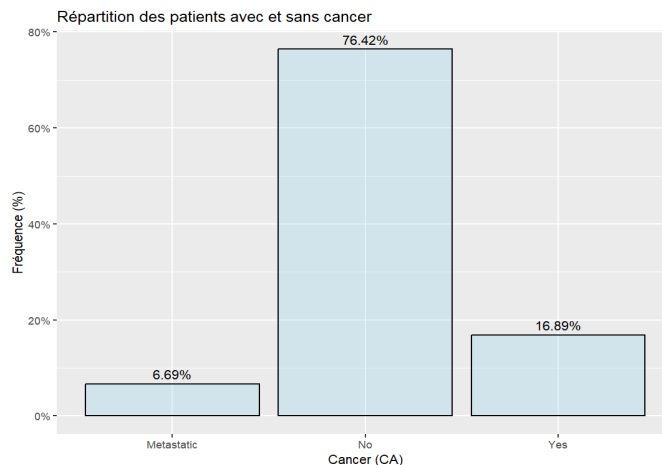
Les caractéristiques sociodémographiques des patients montrent une majorité d'hommes (55.6%) et une prédominance de patients de race blanche (77.45%). La majorité des patients ont suivi entre 10 et 15 années d'études, et 56% ont un revenu annuel inférieur à 11 000€. Les patients sont majoritairement âgés de plus de 50 ans, avec 76% dans cette tranche d'âge. Ces données démographiques fournissent un contexte important pour l'interprétation des résultats cliniques et l'évaluation des besoins en soins.

En conclusion, cette présentation des données met en évidence la diversité et la complexité des informations collectées, essentielles pour évaluer les facteurs de risque et l'impact des interventions médicales sur les résultats cliniques des patients en soins intensifs.

2.2 Variables utilisées

Dans notre étude, 13 variables ressortent de manière significative en raison de leur importance clinique et de leur impact potentiel sur les résultats des patients. Ces variables sont : CA (Cancer), INCOME (Revenu), SWANG1 (Right Heart Catheterization), HEMA (Affection Hématologique), AGE (Âge), DAS2D3PC (Score DASI), et DEATH (Décès), MEANBP1 (Tension Artérielle Moyenne), HRT1 (Fréquence Cardiaque), CARDIOHX (Antécédents Cardiovasculaires), CHFHX (Antécédents d'Insuffisance Cardiaque), CARD (Maladies Cardiaques), AMIHX (Antécédents d'Infarctus du Myocarde). Voici une présentation plus détaillée de chacune d'entre elles :

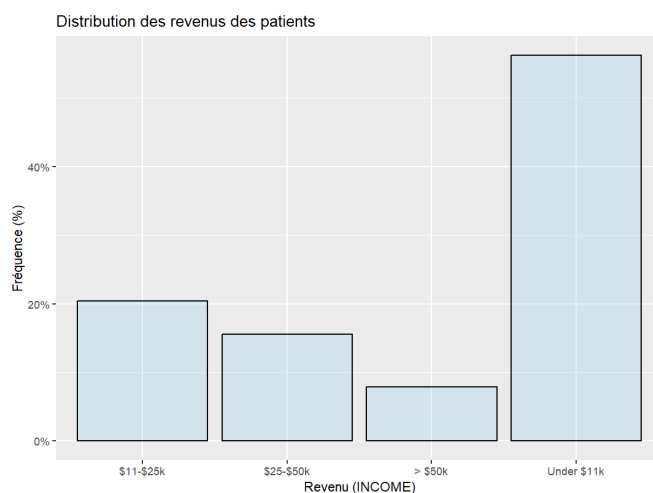
Cancer



1 graphe CA

La variable **CA (Cancer)** indique la présence d'un cancer chez le patient au moment de l'admission en unité de soins intensifs. Cette condition est présente chez 16,89% des patients étudiés, soulignant l'importance de cette comorbidité dans la population analysée. La détection de cancers chez les patients admis en soins intensifs a des implications majeures pour leur traitement et leur pronostic.

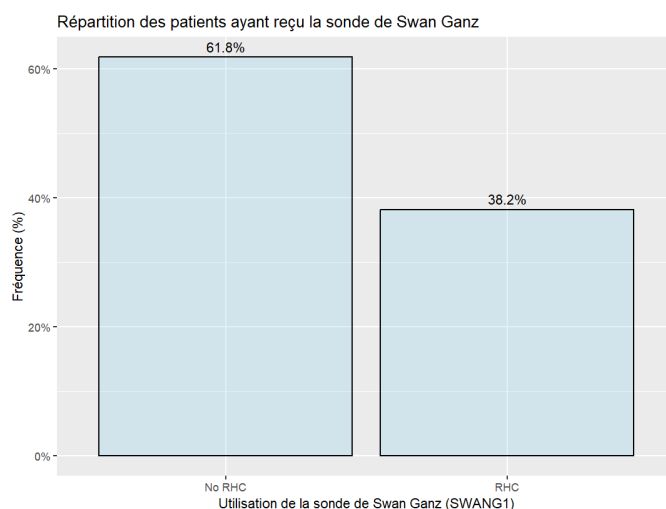
Revenu annuel



2 graphe INCOME

Le **INCOME (revenu annuel)** des patients est une variable sociodémographique cruciale qui peut influencer l'accès aux soins et les résultats de santé. Dans notre étude, 56,18% des patients ont un revenu inférieur à 11 000\$, ce qui pourrait refléter des inégalités dans l'accès aux soins et les résultats cliniques. Les disparités économiques jouent un rôle important dans la santé des patients et leur capacité à recevoir des soins adéquats.

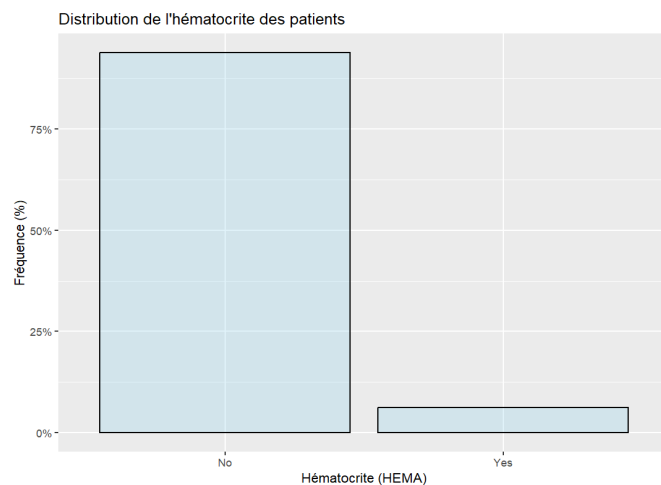
Cathétérissassions cardiaque droite



3 graphe RHC

La variable **SWANG1 (Right Heart Catheterization)** indique si le patient a subi une cathétérissassions cardiaque droite, une procédure utilisée pour diagnostiquer et gérer diverses conditions cardiaques. Cette intervention a été réalisée chez 38,2% des patients, montrant son usage fréquent en soins intensifs malgré les débats sur ses bénéfices versus risques. L'usage de cette technique reste un sujet de discussion quant à son efficacité et ses potentielles complications.

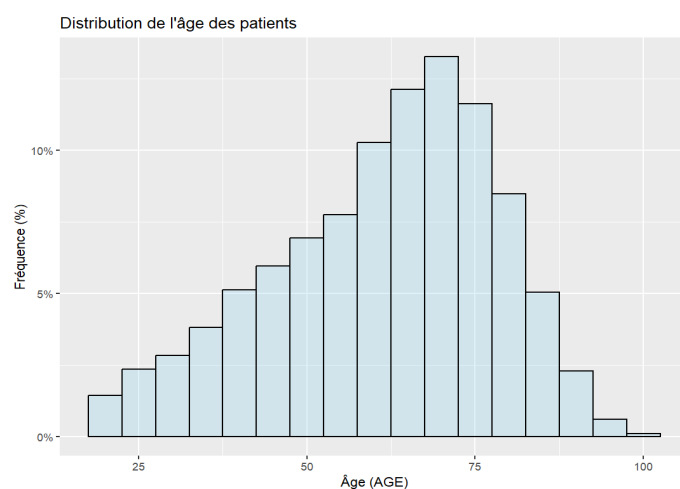
Affections hématologiques



4 graphe HEMA

Les **HEMA** (affections **hématologiques**) , qui incluent des troubles de la coagulation, l'anémie, et d'autres anomalies sanguines, sont représentées par cette variable. 6,25% des patients présentaient une affection hématologique, soulignant l'importance de ces conditions dans la gestion globale des patients en soins intensifs. Les troubles hématologiques peuvent compliquer le traitement et affecter significativement le pronostic des patients.

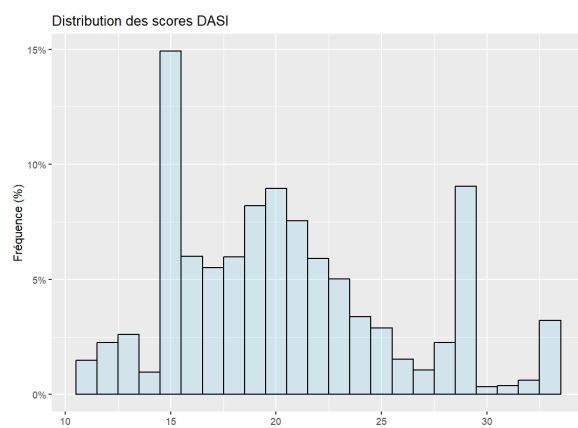
Age



5 graphe AGE

L'**AGE** (**âge**) des patients au moment de leur admission est un facteur crucial influençant la gravité de la maladie et les résultats cliniques. Dans notre étude, 75,27% des patients étaient âgés de plus de 50 ans, ce qui est représentatif de la population vieillissante dans les unités de soins intensifs. L'âge avancé est souvent associé à des comorbidités accrues et à une récupération plus difficile après des interventions médicales intensives.

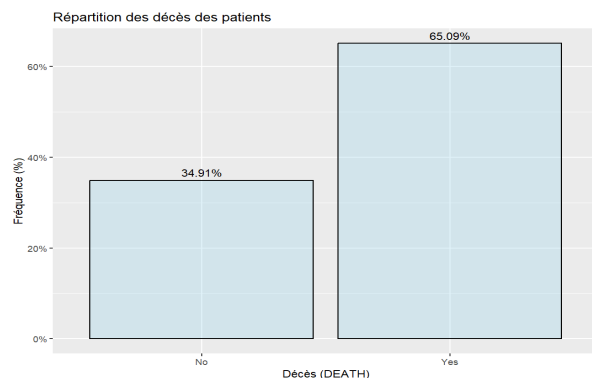
Score DASI



6 graphe DASI

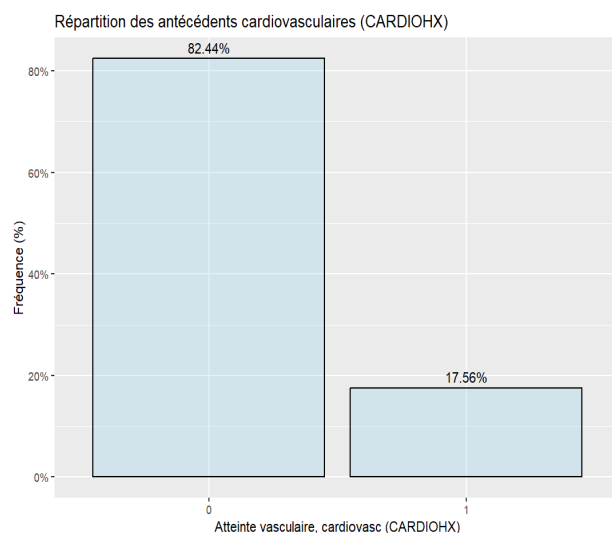
Le **DAS2D3PC** (score **DASI**) évalue la capacité fonctionnelle des patients, avec des scores plus élevés indiquant une meilleure capacité physique. Le score moyen dans notre population est de 20,83, reflétant la variabilité de la capacité fonctionnelle des patients admis en soins intensifs. Ce score est un indicateur important pour évaluer l'état général des patients et leur capacité à mener des activités quotidiennes.

Décès



Enfin, la variable **DEATH (Décès)** indique si le patient est décédé durant la période de suivi de l'étude. 65,05% des patients sont décédés, soulignant la gravité des conditions traitées et la mortalité élevée associée aux soins intensifs. Ce taux de mortalité est un indicateur crucial de l'efficacité des soins et des interventions médicales prodiguées en unité de soins intensifs.

7 graphe DEATH

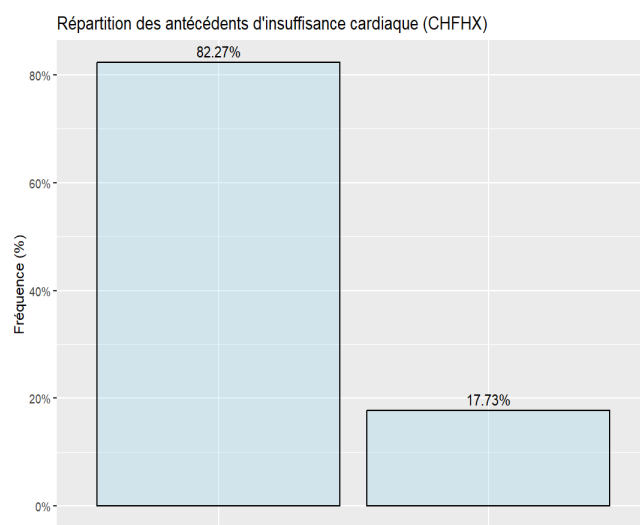


Antécédents vasculaire, cardiovasculaires

La variable **CARDIOHX (antécédents vasculaire, cardiovasculaires)** est importante car elle indique la présence de maladies vasculaires ou cardiovasculaires préexistantes chez les patients admis en USI. Ces antécédents sont présents chez 17,56% des patients étudiés, influençant ainsi leur prise en charge et leur pronostic. Les patients avec des antécédents vasculaires ou cardiovasculaires nécessitent souvent une attention particulière en raison du risque accru de complications.

8 graphe CARDIOHX

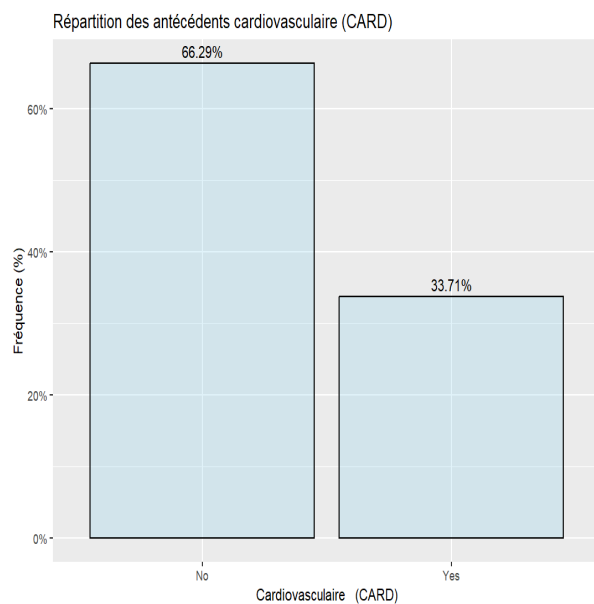
Antécédents d'insuffisance cardiaque



La variable **CHFHX (antécédents d'insuffisance cardiaque)** désigne la présence d'une insuffisance cardiaque dans l'histoire médicale des patients. L'insuffisance cardiaque est une comorbidité sérieuse qui affecte la gestion des patients en USI. Dans notre étude, 17,73% des patients ont des antécédents d'insuffisance cardiaque. La détection de cette condition aide à ajuster les traitements pour stabiliser la fonction cardiaque et prévenir de nouvelles décompensations.

9 graphe CHFHX

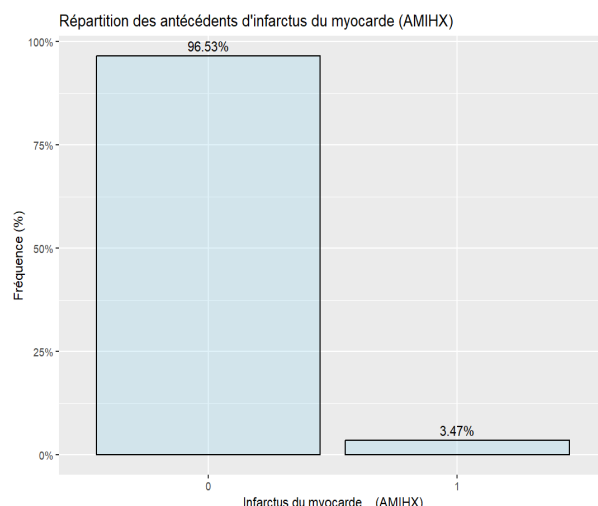
Présence de maladies cardiaques



10 graphe CARD

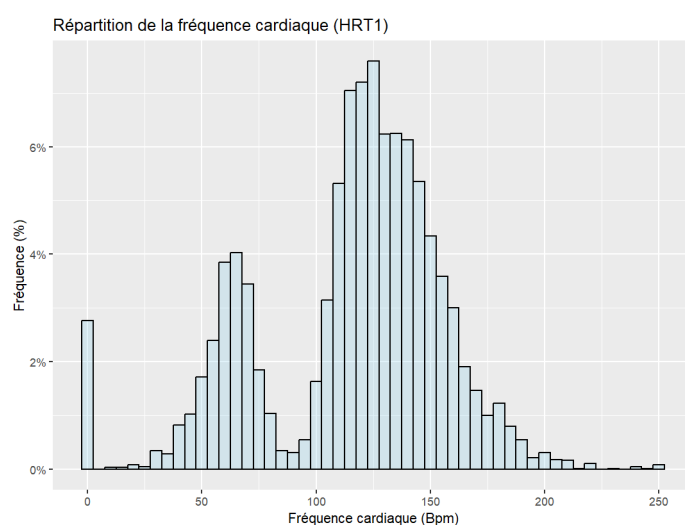
La variable **CARD (présence de maladies cardiaques)** est essentielle pour évaluer l'état de santé cardiovasculaire des patients admis en unité de soins intensifs (USI). La présence de maladies cardiaques chez les patients est une indication critique qui peut influencer le traitement et le pronostic. Dans notre étude, 33,71% des patients présentent des maladies cardiaques, soulignant l'importance de cette comorbidité. La détection et la gestion des maladies cardiaques sont cruciales pour améliorer les résultats cliniques et assurer une prise en charge optimale en soins intensifs.

Antécédents d'infarctus du myocarde



11 graphe AMIHX

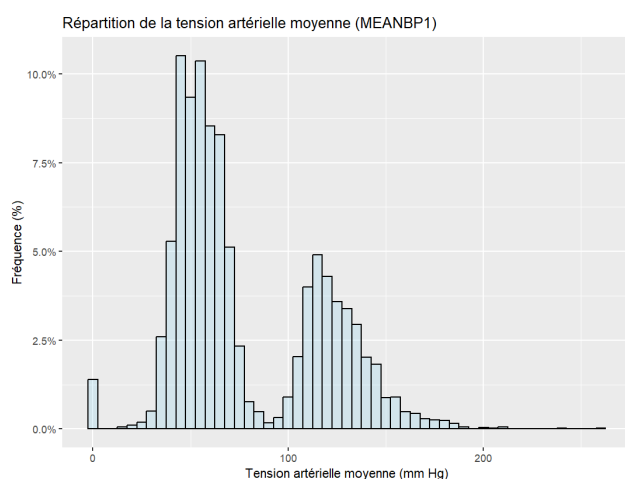
La variable **AMIHX (antécédents d'infarctus du myocarde)** indique si le patient a déjà subi un infarctus du myocarde avant son admission en USI. Les antécédents d'infarctus augmentent le risque de nouveaux événements cardiaques et compliquent souvent le traitement. Dans notre étude, 3,14% des patients ont des antécédents d'infarctus du myocarde. Cette information est cruciale pour la surveillance et la gestion proactive des risques cardiaques pendant le séjour en soins intensifs.



12 graphe HRT1

La variable **HRT1 (fréquence cardiaque)** indique le nombre de battements cardiaques par minute chez les patients admis en USI. Une fréquence cardiaque élevée ou basse peut indiquer une variété de conditions cliniques urgentes. Dans notre étude, 84,19% des patients présentent une tachycardie ou une bradycardie, soulignant l'importance de surveiller et de gérer la fréquence cardiaque pour prévenir les complications graves.

Tension artérielle moyenne



13 graphe MEANBP1

La variable **MEANBP1 (tension artérielle moyenne)** mesure la pression artérielle moyenne des patients admis en unité de soins intensifs (USI). Cette mesure est cruciale pour évaluer l'état hémodynamique des patients et guider les interventions cliniques. Dans notre étude, 45,68% des patients ont une tension artérielle moyenne indiquant un état de choc ou une hypertension sévère, ce qui nécessite une gestion précise pour stabiliser l'état du patient.

Conclusion

Ces treize variables fournissent une vue d'ensemble claire et concise des éléments clés étudiés dans notre population de patients, mettant en lumière leur impact potentiel sur les résultats de santé et la gestion clinique en unité de soins intensifs.

3. Analyse avancée : présentation une par une des problématiques et attentes déclinées dans l'introduction

3.1 Impact des variables sur DEATH

3.1.1 Méthodologie

L'un des objectifs fixé est de mesurer l'impact des différentes variables sur la mort. De cette manière, nous aurions pu déterminer les variables les plus importantes sur lesquelles réaliser des analyses approfondies.

Après des analyses approfondies de ces variables, nous pourrions déterminer l'impact de chacune de ces variables et affirmer que la méthode RHC a un impact significatif sur le fait de décéder par rapport aux autres variables.

Pour sélectionner les variables que nous souhaitons approfondir, nous allons réaliser, en fonction de la nature des variables, des tests du χ^2 ou des tests de comparaison de moyenne, pour déterminer si celles-ci semblent avoir un impact sur la variable DEATH. Nous sélectionnerons ensuite celles qui ont les p-valeurs les plus faibles, autrement dit, dont le lien avec DEATH est relativement certain.

Dans une analyse approfondie, nous allons réaliser des graphiques et étudier plus en profondeur la nature du lien avec la variable DEATH.

3.1.1.1 Test du χ^2

Le test du χ^2 est utilisé pour les variables qualitatives afin de vérifier l'indépendance entre deux variables. Grâce à ce test, nous comparons les distributions observées et attendues des variables pour déterminer s'il existe une association significative entre elles.

Fixons le seuil de significativité à $\alpha = 0,05$, un seuil standard. Le test du χ^2 et le degré de liberté permettent de déterminer la valeur p, ce qui est automatiquement calculé par R.

Sur R, nous utiliserons la commande `chisq.test()`

Voici un exemple de code R :

```
chisq.test(table(donnees_DEATH_modif$CARDIOHX,  
                donnees_DEATH_modif$DEATH))
```

Ce code permet de déterminer directement la valeur du χ^2 , le degré de liberté ainsi que la p-valeur. Nous pouvons ensuite interpréter la p-valeur.

Si la p-valeur est inférieure à α : il existe une association significative.

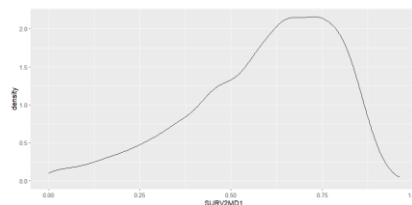
Si P valeur est supérieure α : Nous pouvons conclure que les deux variables étudiées sont statistiquement associées.

3.1.1.2 Test de Wilcoxon

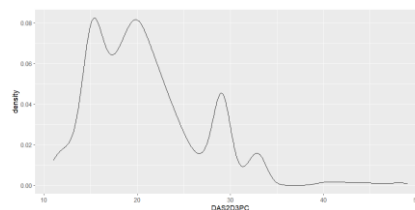
Le test du χ^2 ne s'applique que sur des variables qualitatives. Pour les variables quantitatives, nous allons devoir utiliser un autre type de tests. Des tests de comparaison de moyenne.

Nous avons, dans un premier temps, réaliser des tests de Welch. Néanmoins, comme pour les tests de Student, les tests de Welch nécessitent d'avoir des variables qui suivent, globalement, une distribution normale.

Si certaines variables comme SURV2MD1 peuvent avoir une distribution relativement normale, ce n'est pas le cas d'autres variables comme DAS2D3PC



14 14 - Distribution de la variable SURV2MD1



15 15 - Distribution de la variable DAS2D3PC

Pour un soucis de cohérence, nous décidons d'appliquer un test non paramétrique sur toutes les variables quantitatives. Ce test sera un test de Wilcoxon.

Le test de Wilcoxon, est une méthode statistique non paramétrique utilisée pour déterminer si deux échantillons diffèrent de manière significative.

Sur R, nous utiliserons la commande `wilcox.test()` après avoir fait les éventuels tris que nous avons indiqué lors du rapport intermédiaire informatique concernant les valeurs aberrantes.

```
donnees_temp <- donnees %>% filter(DAS2D3PC<=33)
result <- wilcox.test(donnees_temp[, "DAS2D3PC"]~donnees_temp$DEATH)
```

16 - Exemple de code R pour un test de Wilcoxon

A la suite de cela, nous obtenons une p-valeur qui nous pouvons interpréter. Une faible p-valeur indique une différence significative entre les échantillons de personnes décédées et de personnes non décédées. A partir de cette différence significative, on affirme qu'il existe une relation entre cette variable et DEATH, et que cela vaut la peine d'approfondir cette étude.

Seuls les variables HRT1 (73.03%), PAFI1 (46.2%) et RESP1 (21.5%) ne semblent pas avoir de lien significatif avec la mort. Les autres ont toutes des p-valeur inférieures à 1%, et mis à part pour la variable EDU (0.1%), toutes sont très proches de 0.

Pour sélectionner les variables souhaitées, nous prendrons les extrêmes avec SURV2MD1 (1.495776e-153) et l'AGE (5.428665e-51) et des variables qui nous semblent intéressantes comme DAS2D3PC (1.125829e-30). SURV2MD1 ne sera utilisée que dans la dernière partie.

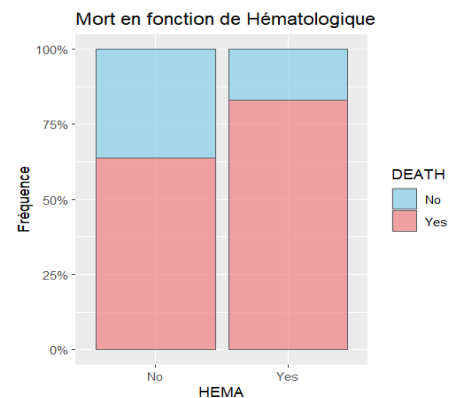
Donc les variables que nous allons sélectionner sont donc : HEMA ; INCOME ; CA ; RHC ; DAS2D3PC ; AGE

3.1.2 Qualitatives

Hématologique

En comparant avec la variable HEMA, on observe une grande différence de proportions, avec un écart significatif de mortalité entre la population ayant reçu HEMA et celle n'en ayant pas reçu. Les personnes n'ayant pas reçu HEMA ont une probabilité de survie plus élevée par rapport à celles qui en ont reçu. Ainsi, il est évident que cette maladie est un facteur important de mortalité.

D'après le test de χ^2 , la p-valeur est égale à $6,262 \times 10^{-13}$, ce qui est inférieur à 0,05. Donc, cette disparité suggère un lien entre la HEMA et une réduction du risque de mortalité.

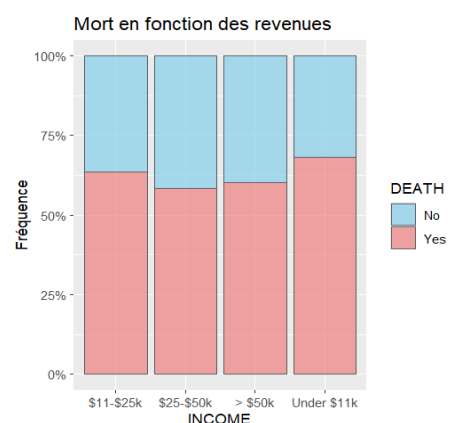


16 graphe DEATH en fonction de HEMA

Revenu

Nous allons maintenant examiner si le revenu a un impact sur la cause de mortalité, ce qui nous permettra d'étudier différentes classes de population. Est-ce que leur décès est influencé par leur niveau de revenu ? Nous pouvons constater que plus la population est pauvre, plus le taux de mortalité est élevé. En revanche, les groupes de revenu entre 25k-50k et au-dessus de 50k ont des taux de mortalité presque identiques.

la p-valeur d'ici est égale à $9,593 \times 10^{-8}$, toujours comparer à 0,05. Cette disparité suggère un lien entre le revenu et une réduction du risque de mortalité

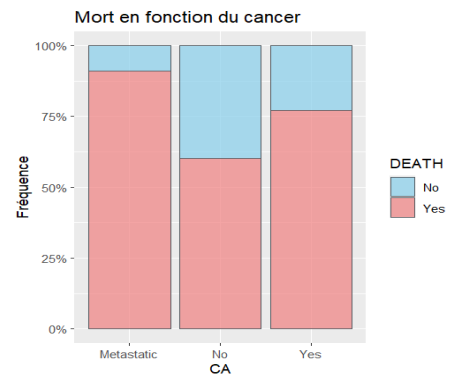


17 graphe DEATH en fonction de INCOME

Cancer

Nous analysons maintenant la cause de mortalité liée au cancer, qui s'avère également très significative. En examinant les détails, nous constatons que les taux de mortalité associés au cancer métastatique et au cancer classique sont très élevés par rapport à ceux des personnes n'ayant pas de cancer (CA).

la p-valeur est égale à $2,2 \times 10^{-16}$, toujours comparer à 0,05. Cette disparité suggère un lien entre le revenu et une réduction du risque de mortalité

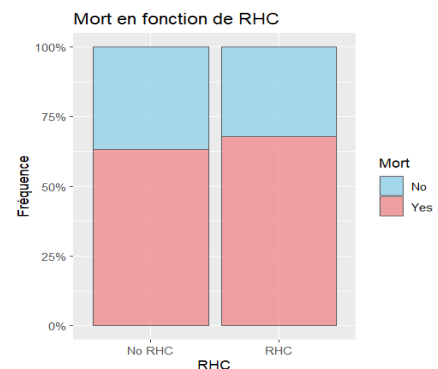


18 graphe DEATH en fonction de CA

RHC

Le diagramme empilé à 100% montre la relation entre la variable RHC et la mortalité. Les données indiquent que parmi les individus non soumis à la RHC, 36,7% ont succombé tandis que 63,3% ont survécu. En revanche, parmi ceux ayant subi la RHC, la proportion de décès s'élève à 32%, tandis que 68% ont survécu. Cela présente une différence de 4,7%.

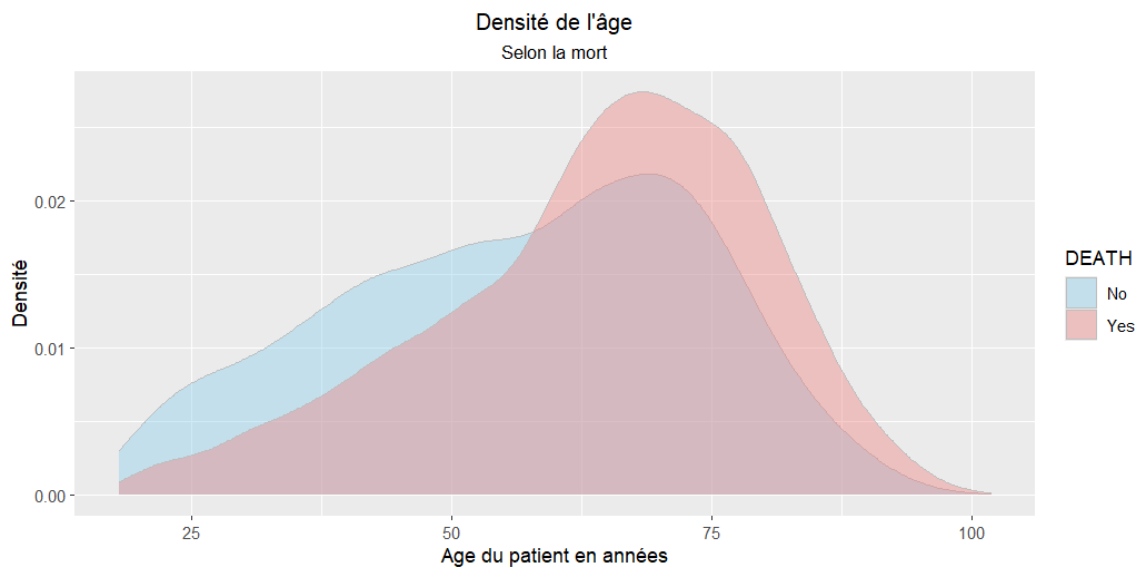
D'après le test de Chi-2, la p-valeur est égale à 0.000388, ce qui est inférieur au seuil de signification de 0.05. Donc, cette disparité suggère un lien entre la RHC et une réduction du risque de mortalité, bien que légère.



19 graphe DEATH en fonction de RHC

3.1.3 Quantitatives

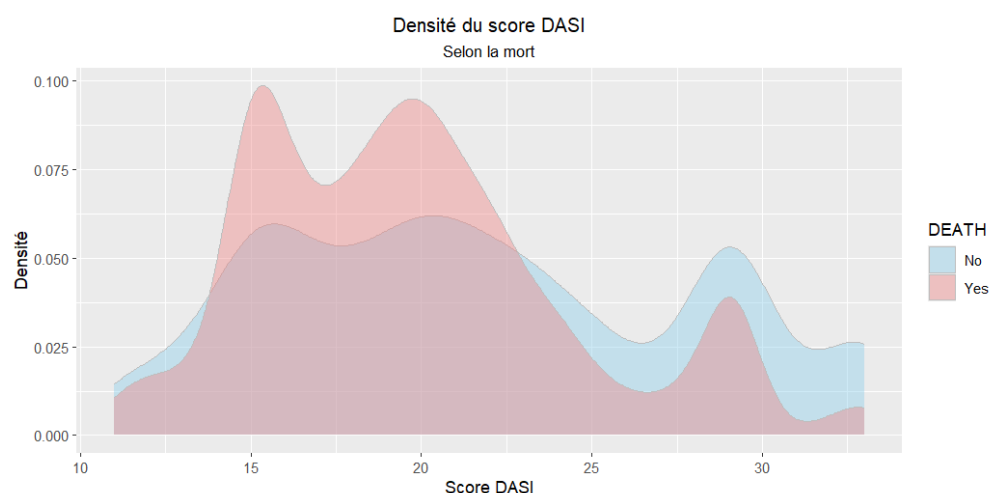
Age



20 densité de l'âge en fonction de DEATH

L'analyse des courbes de densité pour l'âge (AGE) et la mortalité (DEATH) révèle des différences notables entre les patients décédés et survivants. Les courbes montrent que la densité des patients décédés augmente avec l'âge. En particulier, une densité plus élevée de décès est observée chez les patients âgés de plus de 55 ans, suggérant que l'âge avancé est associé à un risque accru de mortalité. Cette tendance peut être attribuée à des facteurs tels que des comorbidités plus fréquentes et une capacité réduite à se remettre de maladies graves chez les personnes âgées. En comparaison, les survivants présentent une distribution d'âge plus uniforme, ce qui indique que les jeunes patients ont une meilleure probabilité de survie. Cette analyse met en évidence l'importance de l'âge comme facteur de risque significatif de mortalité en soins intensifs.

Score DASI



21 Densité du score DASI en fonction de la mort

Les courbes de densité pour le score DASI (DAS2D3PC) et la mortalité (DEATH) montrent une relation inverse entre la capacité fonctionnelle et le risque de décès. Les patients avec des scores DASI inférieurs, indiquant une capacité fonctionnelle réduite, présentent une densité de décès plus élevée, particulièrement concentrée autour des scores de 10 à 22,5. Cela suggère que les patients ayant une capacité fonctionnelle réduite sont plus vulnérables et ont une probabilité plus élevée de décéder. En revanche, les patients survivants tendent à avoir des scores DASI plus élevés, souvent supérieurs à 26, ce qui indique que de meilleures capacités fonctionnelles sont associées à une survie accrue. Cette relation souligne l'importance de la capacité fonctionnelle comme indicateur prédictif de la mortalité en unité de soins intensifs, suggérant que des interventions visant à améliorer la capacité fonctionnelle pourraient potentiellement réduire les taux de mortalité.

3.1.4 Conclusion

Après une analyse approfondie des différentes variables influençant la mortalité, nous avons identifié plusieurs facteurs significatifs dont les impacts varient en intensité.

L'âge (AGE) est un des facteurs les plus influents avec une p-valeur de 5.428665e-51. Cette variable montre que l'âge avancé est fortement associé à une augmentation de la mortalité, suggérant un effet significatif sur les risques de décès.

Le score clinique DASI (DAS2D3PC) présente un test de Wilcoxon de 1.125829e-30 démontrant que les scores inférieurs sont associés à une mortalité plus élevée.

Le revenu (INCOME) avec un test de Wilcoxon de 9.593e-08 indique également une relation significative entre la sécurité financière et la mortalité. Les conditions socioéconomiques défavorisées augmentent le risque de décès.

Le score HEMA, ayant une p-valeur de 6.264e-13, montre une forte association avec la mortalité, soulignant l'importance des paramètres sanguins.

La variable CA, avec une p-valeur de $2.2e-16$ est fortement associée à la mortalité, montrant l'importance des conditions cardiaques.

Enfin, l'impact de la procédure médicale SWANG1 (RHC) sur la mortalité a révélé des résultats inattendus. Bien que l'on puisse anticiper un effet bénéfique, les données montrent que les patients ayant subi la RHC présentent un taux de mortalité plus élevé (32%) comparé à ceux qui n'ont pas eu recours à cette procédure (36.7%). Cette différence de 4.7% en termes de mortalité entre les deux groupes, confirmée par une p-valeur significative de 0.000388, suggère un impact négatif de la RHC sur la mortalité, contrairement aux attentes initiales.

En comparant ces variables entre elles, nous constatons que l'âge (AGE) a l'impact le plus significatif sur la mortalité, suivi de près par les scores DASI et HEMA, puis les conditions socioéconomiques comme le revenu (INCOME). Les conditions cardiaques (CA) et la procédure RHC montrent également des impacts importants, bien que négatifs pour la RHC.

Ainsi, cette étude met en lumière l'importance de divers facteurs dans la mortalité, tout en soulignant les effets potentiellement délétères de certaines interventions médicales, comme la RHC, qui nécessitent une évaluation plus approfondie et des approches alternatives pour améliorer les résultats des patients.

3.2 Impact de la technique RHC sur des variables spécifiques

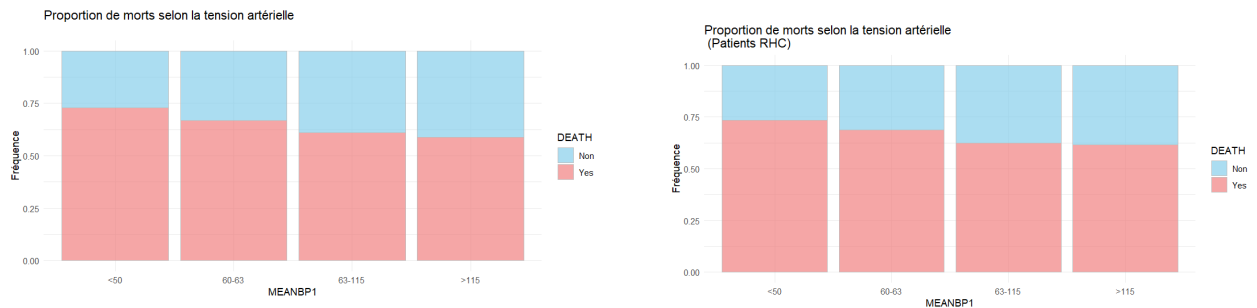
3.2.1 Méthodologie

Comme mentionné précédemment, d'un point de vue global, la RHC a un impact significatif sur la mortalité, ce qui est contraire à nos attentes, car elle augmente le risque de décès. Cela nous a surpris, surtout que nous l'avons découvert à la fin de notre projet. Pour mieux comprendre ces résultats, nous avons donc décidé d'analyser plus en profondeur l'impact de la RHC sur la mortalité en examinant ce même impact dans des populations plus petites. Cependant, par manque de temps, nous n'avons pu effectuer qu'une analyse graphique sans tests statistiques.

Pour sélectionner nos différentes populations, nous avons pris en compte l'ensemble des variables liées au cœur du patient, car, comme mentionné précédemment, la RHC est une technique médicale utilisée pour analyser le cœur. Nous pouvons alors émettre l'hypothèse que RHC ne marche pas sur l'ensemble de la population car ils ne sont pas tous atteints de pathologie cardiaque. Les variables choisis sont la tension artérielle (MEANBP), la fréquence cardiaque (HRT1), l'infarctus du myocarde (AMIHX), les atteintes vasculaires ou cardiovasculaires (CARDIOHX), les maladies cardiovasculaires (CARD), et les crises cardiaques (CHFHX).

3.2.2 Variables

Tension artérielle



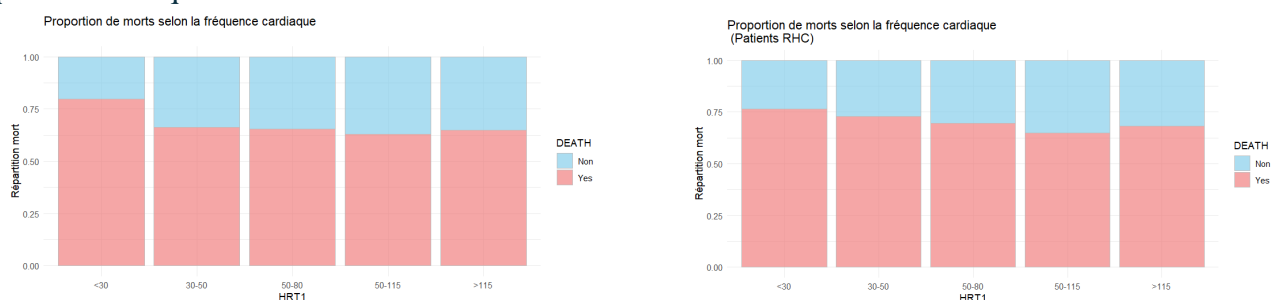
22 graphe DEATH en fonction de MEANBP1

23 graphe RHC en fonction de MEANBP1

La tension artérielle étant une variable continue, nous avons choisi de la regrouper par classes en fonction de ses quartiles.

Sur le graphique de droite, on voit le taux de mortalité pour les différents taux de tension artérielle. Lorsqu'on compare ces différents taux avec ceux du graphique de gauche, on s'aperçoit qu'il n'y a pas d'impact positif sur le taux de survie en fonction d'une tension forte ou d'une tension faible. On retrouve les conclusions sur la population globale qui nous fait dire que les patients qui ont bénéficiés de la RHC meurent plus. Cela peut s'expliquer par le fait qu'une forte ou faible tension n'est pas une pathologie.

Fréquence cardiaque



24 graphe DEATH en fonction de HRT1

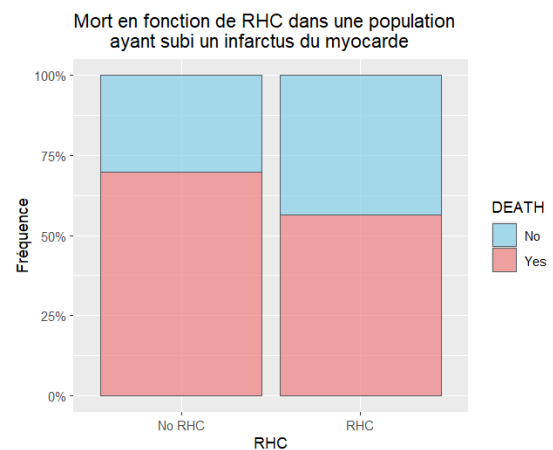
25 RHC en fonction de HRT1

La fréquence cardiaque est une variable qui n'avait pas de fort lien avec la mort, elle a tout de même un lien avec le cœur. HRT1 a été classé en fonction des valeurs « normales » de la fréquence cardiaque. Entre 50 et 80, on a une fréquence classique, au dessus et en dessous, elle commence à devenir anormale.

En comparant les deux graphiques, un impact positif ne peut pas non plus réellement être dégagé. Si sur les fréquence très faible (<30), les patients RHC meurent moins et la RHC serait efficace sur ces classes. Ce n'est pas le cas pour les autres classes. Cela s'explique également par le fait que la fréquence cardiaque n'est pas une pathologie.

Infarctus du myocarde

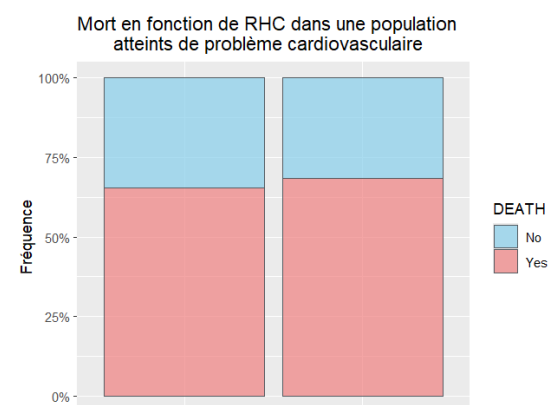
L'histogramme empilé à 100 % illustrant la mortalité en fonction de la RHC chez les patients ayant subi un infarctus du myocarde révèle une différence dans les taux de mortalité entre ceux qui ont bénéficié d'une RHC et ceux qui ne l'ont pas reçue. Dans le groupe sans RHC, environ 70 % des patients sont décédés. En revanche, dans le groupe ayant reçu la RHC, environ 56 % des patients sont décédés, montrant une réduction notable de 14 % du taux de mortalité. Visuellement, cette différence est immédiatement apparente, avec la barre représentant le groupe sans RHC ayant une proportion beaucoup plus élevée de décès par rapport à la barre du groupe avec RHC. Cette réduction du taux de mortalité chez les patients ayant reçu la RHC suggère que cette technique peut effectivement améliorer les chances de survie des patients post-infarctus. En résumé, l'histogramme empilé à 100 % met en évidence l'efficacité apparente de la RHC en réduisant significativement la mortalité, ce qui pourrait encourager son utilisation plus large dans la gestion post-infarctus.



26 graphe DEATH en fonction de RHC dans infarctus

Vasculaire ou cardiovasculaire

L'histogramme empilé à 100 % représentant la mortalité en fonction de la RHC chez les patients atteints de problème vasculaire ou cardiovasculaire montre une tendance différente par rapport aux patients ayant subi un infarctus du myocarde. Dans ce cas, environ 67 % des patients n'ayant pas reçu de RHC sont décédés, tandis qu'environ 72 % des patients ayant reçu la RHC sont décédés. Cette différence de 5 % en faveur d'une mortalité plus élevée chez les patients ayant bénéficié de la RHC est visible sur le graphique.

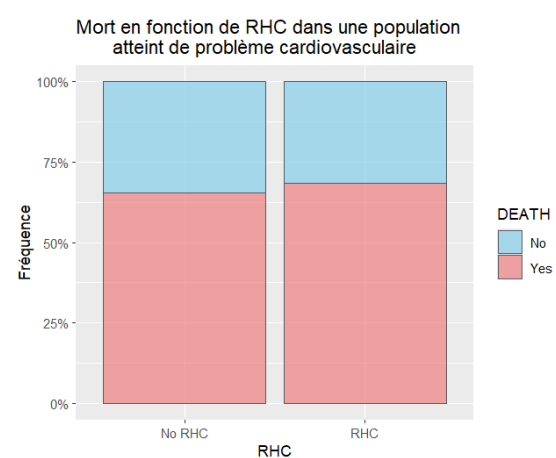


27 graphe DEATH en fonction de RHC dans Cardiovasculaire ou vasculaire

Contrairement aux résultats observés chez les patients post-infarctus, où la RHC semble réduire la mortalité, ici, la RHC est associée à une augmentation du taux de mortalité. Cette tendance pourrait suggérer que la RHC pourrait être néfaste pour les patients souffrant d'atteintes vasculaires ou cardiovasculaires, réduisant ainsi leur taux de survie. Ce contraste met en évidence une non-efficacité, voire un potentiel effet néfaste de la RHC pour ce groupe spécifique de patients. En résumé, l'histogramme empilé à 100 % indique que la RHC comme vue dans la partie précédente, pourrais plutôt que d'améliorer les chances de survie, aggraver la situation des patients avec des atteintes vasculaires ou cardiovasculaires, soulignant ainsi l'importance d'adapter les traitements en fonction des conditions spécifiques des patients.

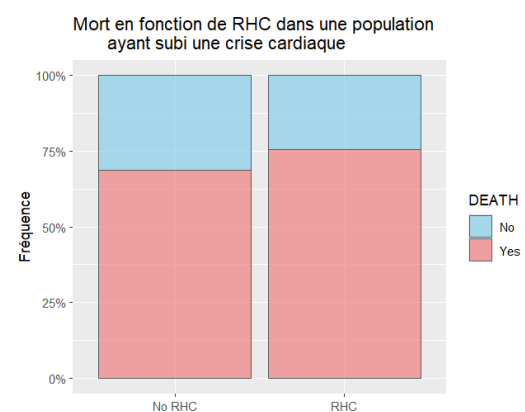
Cardiovasculaire

L'histogramme empilé à 100 % pour les patients ayant une atteinte seulement cardiovasculaire révèle qu'environ 64% des patients n'ayant pas reçu de RHC sont décédés, contre environ 68% des patients ayant reçu la RHC. Bien que la différence soit légère, avec un taux de mortalité 4 % plus haut chez les patients ayant bénéficié de la RHC, cela reste une observation importante. Comme pour les patients ayant une atteinte vasculaire ou cardiovasculaire combinée la RHC semblait augmenter le taux de mortalité.



28 graphe DEATH en fonction de RHC dans cardiovasculaire

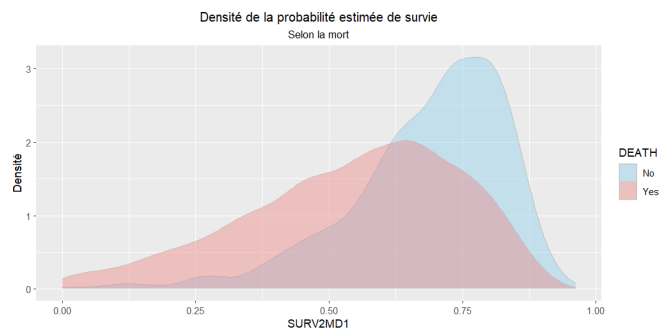
L'histogramme empilé à 100 % pour les patients ayant fait une crise cardiaque révèle qu'environ 69 % des patients n'ayant pas fait de RHC sont décédés, contre environ 75 % des patients ayant bénéficié de la RHC. Cette différence de 6 % en faveur d'une mortalité plus élevée chez les patients ayant reçu la RHC est notable. Comme pour les patients atteints de troubles vasculaires ou cardiovasculaires, on observe une légère différence entre les groupes RHC et non-RHC. Visuellement, la RHC semble augmenter le taux de mortalité,



29 graphe DEATH en fonction de RHC dans crise cardiaque

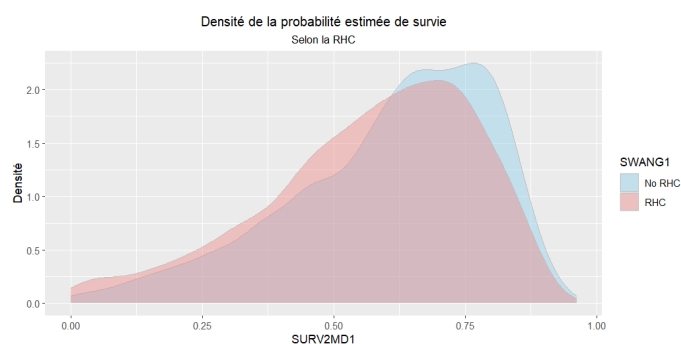
suggérant une réduction de la survie chez les patients ayant subi une crise cardiaque. Cela donne encore une fois l'impression que la technique RHC pourrait être néfaste pour ce groupe spécifique, mettant en évidence une possible inefficacité, voire un effet négatif, de la RHC, similaire à ce qui a été observé dans les premières analyses.

Taux de survie



En regardant le première courbe de densité, on voit, comme nous l'avait affirmés les tests, un lien fortement certains entre le taux de survie estimé et le fait de décéder.

30 courbe de survie selon DEATH



La seconde courbe nous montre qu'il y a également un lien entre ce faible taux de survie et le fait de recevoir le traitement RHC. Les patients qui bénéficient de la RHC ont un pronostic qui leur est moins favorable.

31 courbe de survie selon RHC

A partir de ces deux graphiques, on pourrait émettre l'hypothèse que les patients qui ont bénéficiés de la RHC étaient des patients qui avaient, à l'origine, plus de risques de mourir. Et que la corrélation positive entre la RHC et la mort serait un biais du à cela.

3.2.3 Conclusion

Pour conclure, pour l'infarctus du myocarde, les résultats montrent clairement que la technique RHC a un impact positif sur la mortalité. Les patients ayant subi un infarctus du myocarde et retenu de la RHC ont vu leur taux de mortalité réduit de 14 % par rapport à ceux n'ayant pas reçu cette analyse. Cette réduction notable indique que la RHC améliore les chances de survie dans ce groupe spécifique, pouvant valider l'efficacité de cette technique pour les patients atteints de maladies cardiaques aiguës.

Cependant, pour les patients souffrant de problèmes vasculaires ou cardiovasculaires, de maladies cardiovasculaires, ou ayant subi une crise cardiaque, les résultats ne sont pas aussi prometteurs. En fait, le RHC semble augmenter le taux de mortalité dans ces groupes. Par exemple, les patients avec des atteintes vasculaires ou cardiovasculaires montrent une augmentation de 5 % de la mortalité, et ceux ayant subi une crise cardiaque montrent une augmentation de 6 %. Ces observations suggèrent que la RHC n'est pas aussi efficace, voire potentiellement dangereuse, pour les patients atteints de ces conditions comme vu dans la partie 3.1.

En ce qui concerne le taux de survie global, l'analyse suggère que les patients ayant bénéficié de la RHC avaient initialement des pronostics moins favorables. Cela pourrait confirmer l'hypothèse que les résultats observés sont influencés par la présence de maladies non cardiaques dans la population étudiée. Les courbes de densité montrent un lien entre un faible taux de survie et la réception du traitement RHC, ce qui peut être interprété comme un biais dû au fait que les patients plus à risque de décès ont été ceux principalement traités avec le RHC.

Toutefois, pour les variables comme la tension artérielle et la fréquence cardiaque, le RHC n'a pas montré de différence significative par rapport à la population globale. Les taux de mortalité étaient similaires à la classe de tension artérielle ou à la fréquence cardiaque, ce qui ne soutient pas l'hypothèse que le RHC serait moins efficace sur la population globale en raison de la présence de ces conditions non cardiaques.*

4. Conclusion

Nous avons observé que différentes variables avaient un fort impact, positifs ou négatifs, sur la mort. La gravité du cancer, le fait de souffrir d'une pathologie hématologique à l'admission, un faible revenu, un âge avancé ainsi qu'une mauvaise capacité fonctionnelle vont avoir été significativement lié avec le risque de mourir.

Cependant, bien que la RHC ait un lien significatif avec la mort, le fait de recevoir ce traitement semble également aggraver le risque de survie et non pas le diminuer. La RHC serait donc une technique médicale inefficace, voire dangereuse pour le patient, ce qui nous semble contradictoire.

La RHC étant une technique médicale concernant le cœur, on observe sur les patients qui souffrent de pathologies cardiaques. Bien que la RHC semble améliorer positivement la survie face à l'infarctus du myocarde, c'est la seule des pathologies cardiaques sur laquelle cette technique semble avoir un effet positif significatif. Nous émettons également l'hypothèse que les patients traités étaient, à l'origine, en moins bonne santé que les autres.

Nous répondrons à la problématique posée dans notre introduction en disant que la RHC est uniquement un outil d'évaluation et qu'avec notre base de données datant de la fin du 20^e siècle et que nous ne pouvons pas conclure que la RHC a une influence positive pour les patients.

Notre méthodologie de travail nous a permis d'observer les relations entre les différentes variables en réalisant différents graphiques et en utilisant différents outils statistiques comme le test du χ^2 , les tests de Students et les tests de Wilcoxon. Néanmoins, une trop grande certitude dans les conclusions que nous allions obtenir, nous a amené à être surpris lorsque nous avons réalisé que la technique médicale n'avait pas l'effet escompté. Nous avons dû faire preuve d'adaptabilité pour surmonter cet imprévu.

De plus, certaines pistes explorées, voir réalisées, n'apparaissent pas dans le rapport par manque de place, et ne seront présentes qu'en annexe. Les tests de Students (Welch) n'ont pas été privilégiés au profit des tests de Wilcoxon et un graphique sous forme de heatmap n'a pas été inclus par manque d'adéquation avec nos données. De même, des régressions logistiques étaient prévues mais n'ont pas été réalisées par manque de temps.

Un tableau recensant les résultats de tous nos tests est disponible en annexe.

Ce projet, ces imprévus, les outils statistiques découverts, et notre méthodologie concernant la conduite du projet nous sera utile pour la suite. Comme pistes d'améliorations, nous noterons l'importance de rester méthodique et de toujours vérifier nos postulats de départ.

5. Annexes

1 graphe CA	6
2 graphe INCOME	6
3 graphe RHC	6
4 graphe HEMA	7
5 graphe AGE	7
6 graphe DASI	7
7 graphe DEATH	8
8 graphe CARDIOHX	8
9 graphe CHFHX	8
10 graphe CARD	9
11 graphe AMIHX	9
12 graphe HRT1	10
13 graphe MEANBP1	10
14 14 - Distribution de la variable SURV2MD1	12
15 15 - Distribution de la variable DAS2D3PC	12
16 - Exemple de code R pour un test de Wilcoxon	12
17 graphe DEATH en fonction de INCOME	13
18 graphe DEATH en fonction de CA	14
19 graphe DEATH en fonction de RHC	14
20 densité de l'âge en fonction de DEATH	15
21 Densité du score DASI en fonction de la mort	16
22 graphe DEATH en fonction de MEANBP1	18
23 graphe RHC en fonction de MEANBP1	18
24 graphe DEATH en fonction de HRT1	18
25 RHC en fonction de HRT1	18
26 graphe DEATH en fonction de RHC dans infractus	19
27 graphe DEATH en fonction de RHC dans Cardiovasculaire ou vasculaire	19
28 graphe DEATH en fonction de RHC dans cardiovasculaire	20
29 graphe DEATH en fonction de RHC dans crise cardiaque	20
30 courbe de survie selon DEATH	21
31 courbe de survie selon RHC	21

Codes R des graphiques de la partie 2.3

```
# Fonction pour calculer les pourcentages pour les variables catégorielles
percentage_cat <- function(var) {
  tbl <- prop.table(table(var)) * 100
  data.frame(var = names(tbl), percentage = as.numeric(tbl))
}

# Graphique pour la variable CA
ca_data <- percentage_cat(donnees$CA)
ggplot(ca_data, aes(x = var, y = percentage)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black", alpha = 0.25) +
  geom_text(aes(label = paste0(round(percentge, 2), "%")), vjust = -0.5) +
  labs(title = "Répartition des patients avec et sans cancer",
       x = "Cancer (CA)",
       y = "Fréquence (%)") +
  scale_y_continuous(labels = scales::percent_format(scale = 1))

# Graphique pour la variable INCOME
ggplot(donnees, aes(x = INCOME)) +
  geom_bar(aes(y = (.count..) / sum(.count..) * 100), fill = "skyblue", color = "black", alpha = 0.25) +
  labs(title = "Distribution des revenus des patients",
       x = "Revenu (INCOME)",
       y = "Fréquence (%)") +
  scale_y_continuous(labels = scales::percent_format(scale = 1))

# Graphique pour la variable SWANG1
swang1_data <- percentage_cat(donnees$SWANG1)
ggplot(swang1_data, aes(x = var, y = percentage)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black", alpha = 0.25) +
  geom_text(aes(label = paste0(round(percentge, 2), "%")), vjust = -0.5) +
  labs(title = "Répartition des patients ayant reçu la sonde de Swan Ganz",
       x = "Utilisation de la sonde de Swan Ganz (SWANG1)",
       y = "Fréquence (%)") +
  scale_y_continuous(labels = scales::percent_format(scale = 1))

# Graphique pour la variable HEMA
ggplot(donnees, aes(x = HEMA)) +
  geom_bar(aes(y = (.count..) / sum(.count..) * 100), fill = "skyblue", color = "black", alpha = 0.25) +
  labs(title = "Distribution de l'hématocrite des patients",
       x = "Hématocrite (HEMA)",
       y = "Fréquence (%)") +
  scale_y_continuous(labels = scales::percent_format(scale = 1))
```

Code R des graphiques de la partie 3.1.2

```

# DEATH en fonction de
mes_color <- scale_fill_manual(values = c("Yes" = "lightcoral", "No" = "skyblue"))
|
# RHC
ggplot(quali, aes(x = quali$RHC, fill = factor(quali$DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, color = "#555B61") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction de RHC", x = "RHC", y = "Fréquence", fill = "Mort")

# HEMA
ggplot(quali, aes(x = HEMA, fill = factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, stat = "count", color = "#555B61") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction de Hématologique ", x = "HEMA", y = "Fréquence", fill = "DEATH")

# INCOME
ggplot(quali, aes(x = factor(INCOME), fill = factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, color = "#555B61") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction des revenus", x = "INCOME", y = "Fréquence", fill = "DEATH")

# CA
ggplot(quali, aes(x = CA, fill = factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, stat = "count", color = "#555B61") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction du cancer", x = "CA", y = "Fréquence", fill = "DEATH")

```

Code R de la partie 3.2.2

```
#DEATH en fonction de RHC dans une population

#AMIHX
donnees_AMIHX <- donnees %>%
  filter(donnees$AMIHX == 1)

ggplot(donnees_AMIHX, aes(x = SWANGL, fill = factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, color = "#555861") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction de RHC dans une population
  ayant subi un infarctus du myocarde ", x = "RHC", y = "Fréquence", fill = "DEATH")

#CARDIOHX
donnees_CARDIOHX <- donnees %>%
  filter(donnees$CARDIOHX == 1)

ggplot(donnees_CARDIOHX, aes(x = SWANGL, fill= factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, color = "#555861") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction de RHC dans une population \natteint de problème vasculaire ou cardiovasculaire",
  x = "RHC", y = "Fréquence", fill = "DEATH")

#CARD
donnees_CARD <- donnees %>%
  filter(donnees$CARD == "Yes")

ggplot(donnees_CARD, aes(x = SWANGL, fill = factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, color = "#555861") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction de RHC dans une population
  atteint de problème cardiovasculaire ", x = "RHC", y = "Fréquence", fill = "DEATH")

#CHFHX
donnees_CHFHX <- donnees %>%
  filter(donnees$CHFHX == 1)

ggplot(donnees_CHFHX, aes(x = SWANGL, fill = factor(DEATH))) +
  geom_bar(position = "fill", alpha = 0.7, color = "#555861") +
  mes_color+
  scale_y_continuous(labels = scales::percent_format())+
  labs(title = "Mort en fonction de RHC dans une population
  ayant subi une crise cardiaque", x = "RHC", y = "Fréquence", fill = "DEATH")
```

Tableau récapitulatif des tests .

Variable	Type	Chi²	Wilcoxon	Variable	Type	Chi²	Wilcoxon
PTID	Texte		-	Caractéristiques physiologiques			
Diagnostic à l'admission				WTKILO1	Num.	-	5.128087e-14
CAT1	Texte	2.2e-16	-	TEMP1	Num.	-	2.866504e-12
CAT2	Texte	8.673e-08	-	MEANBP1	Num.	-	2.829464e-13
CA	Texte	2.2e-16	-	RESP1	Num.	-	0.2150184
RESP	Texte	0.07386	-	HRT1	Num.	-	0.7303804
CARD	Texte	0.06607	-	PAFI1	Num.	-	0.4616054
NEURO	Texte	0.8185	-	Caractéristiques sociodémographiques			
GASTR	Texte	0.004286	-	AGE	Num.	.	5.428665e-51
RENAL	Texte	0.1251	-	SEX	Texte	0.2314	.
META	Texte	0.5151	-	RACE	Texte	0.7478	.
HEMA	Texte	6.264e-13	-	EDU	Num.	.	0.001021223
SEPS	Texte	0.005425	-	INCOME	Texte	9.593e-08	-
TRAUMA	Texte	7.323e-06	-	NINSCLAS	Texte	2.2e-16	-
ORTHO	Texte	0.6104	-				
Maladies associées							
CARDIOHX	Num.	0.003995	-				
CHFHX	Num.	4.601e-06	-	Evénements			
DEMENTHX	Num.	5.745e-10	-	SWANG1	Texte	0.0003884	
PSYCHHX	Num.	0.01261	-	DEATH	Texte	-	-
CHRPULHX	Num.	0.02269	-	Dates			
RENALHX	Num.	0.1189	-	date_inclusion _etude	Num.		
LIVERHX	Num.	0.0205	-	date_sortie _hopital	Num.		
GIBLEDHX	Num.	0.1662	-	date_deces	Num.		
MALIGHX	Num.	2.2e-16	-	date_dernière _nouvelle	Num.		
IMMUNHX	Num.	5.61e-05	-	Durées			
TRANSHX	Num.	7.196e-07	-	Duree_sejour	Num.		
AMIHX	Num.	0.6627	-	Duree_survie _inclusion	Num.		
Scores Cliniques				duree_survie _sortie	Num.		
SURV2MD1	Num.	.	1.495776e-153				
DAS2D3PC	Num.	-	1.125829e-30				
APS1	Num.	-	1.340561e-46				
SCOMA1	Num.	-	1.041382e-23				

Code R pour réaliser des tests de Students (à l'aide d'une boucle)

```
##STUDENTS
num_col <- c("AGE", "EDU", "SURV2MD1", "HRT1", "TEMP1", "PAFI1", "WTKILO1")
for (i in num_col) {
  result <- t.test(donnees_DEATH_modif[,i]~donnees_DEATH_modif$DEATH)
  print(cat(i, result$p.value, "\nMoyennes estimées :\n", result$estimate,
           "\nIC95% de l'écart entre les groupes :\n", result$conf.int, end="\n——\n\n"))
}
```

Code heatmap (ne convient pas).

```
maladi <- maladi %>%
  mutate(DEATH = ifelse(DEATH == "Yes", 1, 0))

maladi <- maladi %>%
  mutate(RHC = ifelse(RHC == "RHC", 1, 0))

# Supposons que maladi soit déjà votre dataframe
# Convertir les colonnes catégorielles en facteurs puis en numériques
maladi$SEX <- as.numeric(as.factor(maladi$SEX))
maladi$RACE <- as.numeric(as.factor(maladi$RACE))
maladi$INCOME <- as.numeric(as.factor(maladi$INCOME))
maladi$NINSCLAS <- as.numeric(as.factor(maladi$NINSCLAS))

# Vérifier la structure des données
str(maladi)

# Sélectionner uniquement les colonnes numériques pour la matrice de corrélation
colonnes_numeriques <- sapply(maladi, is.numeric)
maladi_numerique <- maladi[, colonnes_numeriques]

# Calculer la matrice de corrélation
matrice_correlation <- cor(maladi_numerique, use = "complete.obs")

# Transformer la matrice de corrélation en un format adapté pour ggplot2
cor_data <- melt(matrice_correlation)

# Créer la carte thermique
ggplot(data = cor_data, aes(x = Var1, y = Var2, fill = value)) +
  geom_tile() +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
                      midpoint = 0, limit = c(-1, 1), space = "Lab",
                      name="Corrélation") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1,
                                    size = 10, hjust = 1)) +
  coord_fixed()

heatmap(matrice_correlation)
```

