DEVOIR EPIDEMIOLOGIE : UTILISATION DU SCORE DE PROPENSION

Sarah F FELDMAN

27 février 2017

# INTRODUCTION

Le cathétérisme cardiaque droit (CCD) est un examen invasif utilisé en soins intensif pour mesurer directement la fonction cardiaque, ce qui permettrait selon certains médecins une meilleure prise en charge du patient et donc une augmentation de la survie. Cependant pour montrer une telle causalité il faudrait un essai randomisé. Or nous avons ici les données d'une cohorte prospective, non interventionnelle, de patients hostpitalisés en réanimation, certains ayant été cathétérisés, d'autres non. La cathétérisation n'a pas été randomisée, il y a donc lieu de penser que l'examen a été réalisé préférentiellement chez certains types de patients, potentiellement les patients les plus graves qui ont donc plus de risque de décéder. Cela constitue un biais d'indication qui rend l'analyse de l'efficacité du CCD ininterprêtable. Pour pouvoir analyser l'efficacité du CCD par sonde de swan Ganz, nous allons prendre en compte les biais d'indication propables à l'aide d'un score de propension, nous permettant d'améliorer le niveau de causalité de la relation CCD/Décès si elle existe.

# I - ETAPES PRELIMINAIRES

## 1) Description de la base de donnees

La base de données comporte les informations de 5735 patients pour 63 variables. Ce sont des données transversales avec une ligne par patient (pas de doublon), chaque variable ayant été mesurées une seule fois.

## 2) Data management

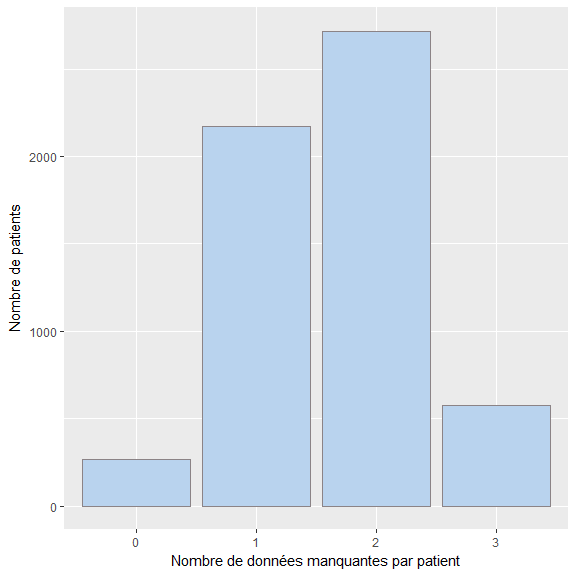
Je fais une première étape de data management "basique" pour mettre les variables au bon format (dates, numérique et facteur), m'assurer que le numero de patient est en caractère, créer une variable SWAN en 0/1 et une autre en TRUE/FALSE selon les différents modèles utilisés par la suite. Je change également les "" en "no category"" pour la variable CAT2 ou en NA pour les dates. Je choisis No comme référence pour la variable CA.

## 3) Vérification de la durée de suivi

Je vérifie les variables dates de décès, date de dernieres nouvelles, et durée de suivi pour l'outcome décès à 30 jours :

* 3 patients pour lesquels la date de décès est antérieure à la date de derniere nouvelle et 2151 pour lesquels date de dernières nouvelles est antérieure au décès. Je modifie la variable date de dernières nouvelles ; si la datee de décès est non nulle, alors la date de dernières nouvelles est la date de décès.
* Nous prendrons comme outcome le décès à 30 jours, je dois donc utiliser les variables DTH30, qui est le décès à 30 jours et T3D30 qui est le temps de suivi adapté à cet outcome. Afin de vérifier ces 2 variables, je recree la variable temps de suivi pour l'outcome décès à 30 jours.  
  Pour cela je cree d'abord la variable temps de suivi comme la différence entre la date des dernières nouvelles et la date d'admission dans l'étude. Puis je modifie cette variable telle que si le temps de suivi est supérieur à 30 jours alors je le ramène à 30 jours. Et la variable décès à 30 jours est modifiée comme ceci à partir de la variable décès : si le décès est survenu après 30 jours, alors la variable décès à 30 jours vaut 0.  
  En comparant les variables DTH30 et T3D30 avec celles recrées, je trouve une différence pour 8 patients qui ont un temps de suivi inférieur à 30 jours et qui sont pourtant noté avec un temps de suivi de 30 jours pour un outcome décès à 30 jours.Et ça ne vient pas de la modification de la variable date de dernières nouvelles. Pour la suite du devoir, je décide de prendre T3D30 comme variable de temps de suivi pour le décès à 30 jours.

## 4) Données manquantes



Figue 1. Nombre de valeurs manquantes par patient.

La base de données est globalement de bonne qualité, en effet aucun sujet n'a plus de 4 valeurs manquantes (sur les 63 variables). (Figure 1)

Si je regarde maintenant le nombre de valeurs manquantes pour chaque variable, je vois que ADLD3P (échelle ADL) et URIN1 (volume urinaire des 24h) ont respectivivement 75% et 53% de données manquantes tandis que les autres variables n'ont aucune ou moins de 5% de données manquantes.

## 5) Analyse descriptive préliminaire afin de repérer les incohérences

### a- Présentation des variables de l'examen clinique et biologique :

* ADLD3P = Activities of daily living (ADL) : échelle d'autonomie de 0 à 12 (score>6 signe une dépendance).
* DAS2D3PC = Duke Activity Status Index (DASI) : auto-questionnaire de 12 items mesurant l'activité fonctionnelle. Le score va de 0 à 58.2, plus le score est élevé, meilleur est l'activité fonctionnelle.
* DNR1 = Do not resuscitate : 1 pour une interdiction de réanimation cardiopulmonaire, 0 sinon.
* SURV2MD1 = Probabilité de survie à 2 mois, estimée par model : de 0 à 1.
* APS1 = APACHE III Acute Physiology scores : score de prediction du risque de mortalité de patients hospitalisés en soins intensifs. Plus le score est élevé plus le risque de mortalité est important. Le score va de 0 à 299 mais ici seule la partie physiologie du score est utilisée.
* SCOMA1 = score de Glasgow. Ce score évalue l'état de conscience du patient: un score de 3 équivaut à un coma profond, un score de 15 est un état de conscience normal.
* PAFI1 = rapport PAO2/FIO2. Il permet de diagnostiquer une agression pulmonaire aigue(rapport>300), un rapport <200 definissant le syndrome de detresse respiratoire aigue et l'ECMO est envisageable en cas de rapport <50.
* PH1 = pH. Un pH normal varie entre 7,38 et 7.42.
* HEMA1 = taux d'hématocrite, normalement compris entre 41 et 50% environ.
* PACO21 = Pression artérielle en CO2. La PaCO2 normale varie de 35 à 45mmHg.
* ALB1 = taux d'albumine. Le taux d'albumine normal varie entre 25 et 44 g/L.
* WTKILO1 = poids
* TEMP1 = température corporelle
* HRT1 = fréquence cardiaque
* MEANBP1 = pression artérielle moyenne
* RESP1 = fréquence respiratoire
* SOD1 = natrémie
* POT1 = kaliémie
* CREA1 = créatininémie
* BILI1 = bilirubinémie
* URIN1 = volume urinaire des 24h

### b- valeurs aberrantes

Pour détecter les valeurs aberrantes je dispose de plusieurs méthodes. Je regarde le tableau descriptif en calculant pour chaque variable quantitative la moyenne, l'intervalle interquartile et surtout le range (tableau non présenté). Je croise également les variables fréquence cardiaque, pression artérielle et fréquence respiratoire à la recherche d'incohérence.

Enfin je regarde la distribution des variables ce qui me permet de réperérer éventuellement d'autres incohérences (voir annexe : Figure A-1). Par exemple nous voyons un pic à 0 pour la variable WTKILO1 (nous l'avions déjà vu grâce au tableau descriptif, mais c'est un deuxième filet de sécurité). On observe également une pause dans les inclusions(variable SADMDTE), mais je ne sais pas comment le prendre en compte.

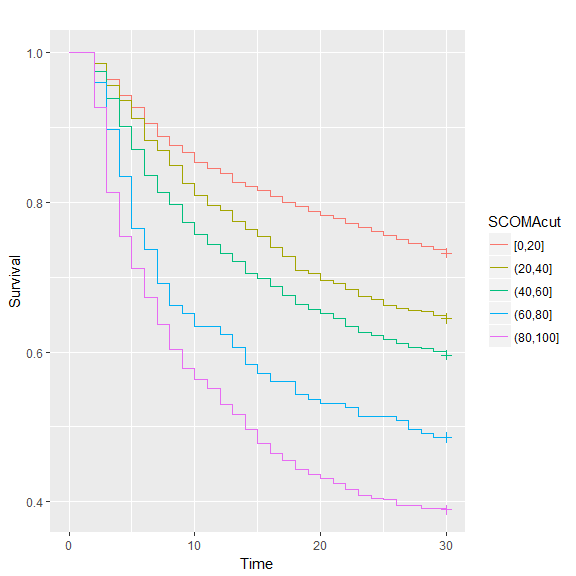


Figure 2. Décès en fonction du score de glasgow découpé en quintile.

Les scores qui ont été mesurés, semblent tous avoir des résultats plausibles excepté le score de Glasgow qui a été modifié : d'un score allant de 3 à 15 on passe à un score allant de 0 à 100. Le score de 100 est probablement lié à un coma sévère car les patients avec un score plus élevé ont un risque de décès augmenté (alors que dans l'échelle d'origine, 3 correspond au coma sévère, et 15 à une conscience normale) (Figure 2).  
Je laisse la variable telle quelle car elle a de toute évidence été modifiée, je ne veux pas la remodifier une seconde fois.  
La température corporelle peut aller de 27° en cas d'hypothermie profonde à 42-43° en cas de forte fièvre, donc je ne modifie pas.  
Un patient ne peut pas avoir un poids de 0, ces 515 patients sont recodés en NA pour le poids.  
Un patient peut être en aplasie mais je ne suis pas sûr qu'une hyperleucocytose de plus de 40 leucocytes/10^9L soit possible, cependant dans le doute concernant le taux maximum en cas de leucémie et au vue de la distribution de la variable qui semble plausible, je laisse telle quelle.  
Hematocrite : une anémie chronique peut entrainer une anémie profonde avec une hématocrite de 10%, en dessous de cette valeur je mets NA. (un taux de 66% est possible). Un taux de PAcO2 de 156 mmHg me semble complètement impossible, au delà de 100 mmHg je met NA. Un pH de 6.6 est extrêmement faible et n'est pas viable mais je ne suis pas sûre qu'on ne puisse pas l'observer en réanimation.  
Une créatininémie de 2200umol/L ou 25 mg/dL est possible en cas d'insuffisance rénale terminale. Un volume urinaire de 0 et de 9000mL est possible également(mais cette variable n'est de toute façon pas utilisée pour l'analyse).  
Une valeur normale d'albumine se situe entre 3.5 et 5 g/dL. Les patients avec un taux d'albumine supérieur à 10g/dL auront une valeur NA pour l'albumine. Les valeurs de natrémie, kaliémie, fréquence respiratoire, fréquence cardiaque, pression artérielle moyenne et bilirubine semblent plausibles.

En observant ensemble la fréquence cardiaque, la fréquence respiratoire et la tension artérielle, je note certaines incohérences : parfois 2 de ces variables peuvent valoir 0 mais pas la troisième, et parfois une seule de ces variables vaut 0. Ainsi 1 patient à une tension artérielle à 0 mais une fréquence cardiaque et une fréquence respiratoire différente de 0, 47 patients ont une fréquence respiratoire à 0 mais une fréquence cardiaque et une tension artérielle différente de 0 et 3 patients ont une fréquence cardiaque à 0 mais une fréquence respiratoire et une tension artérielle différente de 0.

Je modifie donc ces 3 variables selon les règles suivantes :

* Si deux de ces variables valait 0, alors la valeur de 0 était systématiquement attribuée à la troisième.
* Si une seule de ces variables valait 0 alors elle était systématiquement transformée en valeur manquante.

# II - IMPUTATION DES VALEURS MANQUANTES

Les valeurs manquantes doivent absolument être prises en charge pour la suite du projet car on ne peut pas faire de score de propension pour les patients ayant une ou plusieurs variables explicatives manquantes.

## 1) Règles d'imputation

Règles de décision concernant l'imputation ou non des variables :

* une variable avec plus de 20% de données manquantes ne sera pas imputée et ne sera pas inclue dans l'analyse.
* un patient avec plus de 50% de données manquantes ne sera pas inclu dans l'analyse.
* un patient n'ayant pas d'information concernant la variable swan ganz qui est la variable explicative d'interet ne sera pas inclu dans l'analyse car non informatif.
* j'imputerai les variables avec des valeurs manquantes uniquement si plus de 5% des patients on des valeurs manquantes, sinon je supprimerai simplement ces patients.

## 2) Description des valeurs manquantes (valeurs aberrantes transformées en valeurs manquantes)

Je vais donc dans un premier temps observer les données manquantes par variable et par sujet, après avoir transformé les valeurs aberrantes en valeurs manquantes.

Comme vu précédemment, Le score d'autonomie ADL(ADLD3P) a environ 75% de données manquantes et la variable volume d'excrétion urinaire (URIN1) a plus de 50% de données manquantes. Je supprime donc ces deux variables de l'analyse.  
250 patient patients n'ont aucune donnée manquante, 5485(95.6%) patients ont au moins 1 donnée manquante dont : 1930 patients avec 1 donnée manquante, 2698 patients avec 2 de données manquantes, 796 patients avec 3 données manquantes, 64 patients avec 4 données manquantes et 1 patient avec 5 données manquantes. Tous les patients ont donc moins de 50% de données manquantes.  
Il n'y a pas de valeurs manquantes concernant la variable explicative d'intérêt SWAN (SWAN = 0 pas de traitement par CCD, SWAN = 1 traitement par CCD).

En supprimant ADLD3P et URIN1 de l'analyse, j'ai 608 patients avec au moins une donnée manquante, soit plus de 5 % de patients. Je ne peux donc pas simplement supprimer les patients avec des données manquantes. 6 variables ont au moins une donnée manquante : la tension artérielle (15 NA), la fréquence cardiaque (60 NA), l'albuminémie (2 NA) l'hématocrite (8 NA), la PaCO2 (21 NA) et le poids (515 NA).

J'impute les variables explicatives avec le package mice en faisant l'hypothèse que les données manquantes le sont aléatoirement. J'utililse la technique d'imputation multiples mais je ne réalise qu'un seul jeu de de données imputé. Toutes les variables du jeu de données sont prédictrices (y compris l'outcome DTH30, le temps de suivi jusqu'à cet outcome et SWAN, selon Jonathan A C Sterne et al 2009 BMJ) excepté les dates. Par contre je ne prend pas en compte DEATH qui n'est pas notre outcome ni le temps de suivi jusqu'à DEATH.

iter imp variable  
 1 1 MEANBP1 HRT1 ALB1 HEMA1 PACO21 WTKILO1  
 2 1 MEANBP1 HRT1 ALB1 HEMA1 PACO21 WTKILO1  
 3 1 MEANBP1 HRT1 ALB1 HEMA1 PACO21 WTKILO1  
 4 1 MEANBP1 HRT1 ALB1 HEMA1 PACO21 WTKILO1  
 5 1 MEANBP1 HRT1 ALB1 HEMA1 PACO21 WTKILO1

## 3) Descrition des variables imputées avant et après imputation

Je compare les variables qui ont été imputées, avant et après imputation. Je ne compare que visuellement, je ne fais pas de test.

|  |  |  |  |
| --- | --- | --- | --- |
|  | valeur | missing values | range |
| MEANBP1\* | 63 (50-115) | 15 (0.3%) | 0 - 259 |
| HRT1\* | 124 (100-142) | 60 (1%) | 0 - 250 |
| ALB1\* | 3.5 (2.6-3.5) | 2 (0%) | 0.299987793 - 6.599609375 |
| HEMA1\* | 30 (26.1-36.3) | 8 (0.1%) | 10 - 66.1875 |
| PACO21\* | 37 (31-42) | 21 (0.4%) | 1 - 100 |
| WTKILO1\* | 72.2 (60.4-85.2) | 515 (9%) | 19.5 - 244 |

Table 1. Variables qui ont été imputées, avant imputation

|  |  |  |  |
| --- | --- | --- | --- |
|  | valeur | missing values | range |
| MEANBP1\* | 63 (50-114.5) | 0 (0%) | 0 - 259 |
| HRT1\* | 124 (100-142) | 0 (0%) | 0 - 250 |
| ALB1\* | 3.5 (2.6-3.5) | 0 (0%) | 0.299987793 - 6.599609375 |
| HEMA1\* | 30 (26.1-36.3) | 0 (0%) | 10 - 66.1875 |
| PACO21\* | 37 (31-42) | 0 (0%) | 1 - 100 |
| WTKILO1\* | 72.2 (60.4-85.2) | 0 (0%) | 19.5 - 244 |

Table 2. Variables qui ont été imputées, après imputation

La distribution des variables a l'air semblable avant et après imputation.

# III - DESCRIPTION DE LA POPULATION DE L'ETUDE (après imputation)

|  |  |  |  |
| --- | --- | --- | --- |
| SEX\_Female | 1 | 2 | 1 |
| SEX\_Male | 2 | 1 | 2 |
| RACE\_black | 5 | 5 | 5 |
| RACE\_other | 3 | 3 | 3 |
| RACE\_white | 4 | 4 | 4 |
| INCOME\_$11-$25k | 9 | 9 | 6 |
| INCOME\_$25-$50k | 8 | 8 | 9 |
| INCOME\_> $50k | 7 | 7 | 8 |
| INCOME\_Under $11k | 6 | 6 | 7 |
| NINSCLAS\_Medicaid | 12 | 12 | 15 |
| NINSCLAS\_Medicare | 14 | 14 | 11 |
| NINSCLAS\_Medicare & Medicaid | 11 | 10 | 14 |
| NINSCLAS\_No insurance | 10 | 11 | 13 |
| NINSCLAS\_Private | 15 | 15 | 12 |
| NINSCLAS\_Private & Medicare | 13 | 13 | 10 |
|  | 16 | 16 | 16 |
|  | 17 | 17 | 17 |
| CAT1\_ARF | 18 | 25 | 20 |
| CAT1\_CHF | 21 | 20 | 24 |
| CAT1\_Cirrhosis | 19 | 21 | 19 |
| CAT1\_Colon Cancer | 26 | 18 | 26 |
| CAT1\_Coma | 23 | 26 | 23 |
| CAT1\_COPD | 24 | 23 | 25 |
| CAT1\_Lung Cancer | 22 | 22 | 21 |
| CAT1\_MOSF w/Malignancy | 20 | 19 | 22 |
| CAT1\_MOSF w/Sepsis | 25 | 24 | 18 |
| CAT2\_Cirrhosis | 30 | 27 | 30 |
| CAT2\_Colon Cancer | 27 | 18 | 28 |
| CAT2\_Coma | 33 | 30 | 33 |
| CAT2\_Lung Cancer | 28 | 29 | 27 |
| CAT2\_MOSF w/Malignancy | 29 | 23 | 29 |
| CAT2\_MOSF w/Sepsis | 32 | 31 | 32 |
| CAT2\_NoCAT2 | 31 | 28 | 31 |
| CA\_No | 35 | 32 | 35 |
| CA\_Metastatic | 34 | 10 | 34 |
| CA\_Localized | 36 | 33 | 36 |
| CARDIOHX\_1 | 38 | 35 | 37 |
| CHFHX\_1 | 40 | 37 | 39 |
| DEMENTHX\_1 | 42 | 38 | 42 |
| PSYCHHX\_1 | 43 | 40 | 43 |
| CHRPULHX\_1 | 46 | 43 | 45 |
| RENALHX\_1 | 47 | 44 | 47 |
| LIVERHX\_1 | 49 | 11 | 49 |
| GIBLEDHX\_1 | 51 | 48 | 51 |
| MALIGHX\_1 | 54 | 50 | 53 |
| IMMUNHX\_1 | 56 | 52 | 55 |
| TRANSHX\_1 | 58 | 54 | 58 |
| AMIHX\_1 | 59 | 26 | 59 |
|  | 61 | 56 | 61 |
| DNR1\_1 | 63 | 57 | 63 |
|  | 64 | 59 | 64 |
|  | 65 | 60 | 65 |
|  | 66 | 61 | 66 |
|  | 67 | 62 | 67 |
|  | 68 | 63 | 68 |
|  | 69 | 64 | 69 |
|  | 70 | 65 | 70 |
|  | 71 | 66 | 71 |
|  | 72 | 67 | 72 |
|  | 73 | 68 | 73 |
|  | 74 | 69 | 74 |
|  | 75 | 70 | 75 |
|  | 76 | 71 | 76 |
|  | 77 | 72 | 77 |
|  | 78 | 73 | 78 |
|  | 79 | 74 | 79 |
|  | 80 | 75 | 80 |
|  | 81 | 76 | 81 |
| RESP\_1 | 82 | 78 | 82 |
| CARD\_1 | 84 | 80 | 84 |
| NEURO\_1 | 87 | 81 | 87 |
| GASTR\_1 | 89 | 31 | 89 |
| RENAL\_1 | 90 | 84 | 90 |
| META\_1 | 92 | 87 | 92 |
| HEMA\_1 | 94 | 88 | 94 |
| SEPS\_1 | 97 | 91 | 96 |
| TRAUMA\_1 | 98 | 93 | 98 |
| ORTHO\_1 | 100 | 95 | 26 |

Table 3. Caractéristiques de base des patients (après imputation des valeurs manquantes).  
\*variable quantitative

Il me manque une information essentielle : je ne sais pas si les variables ont été mesurée à l'admission ou après pose de la sonde de swan ganz. Je vais considérer que c'est avant pose de la sonde, sinon ça n'a pas de sens.

Je regarde les différences de caractéristiques de base entre le groupe des patients cathétérisés et le groupe des patients non cathétérisés (Table 3). Je ne réalise pas de test car une petite différence numériquement peut être significative du fait de la taille de l'échantillon.  
Les patients traités par CCD sont plus souvent des hommes, avec une couverture médicale moins précaire, ils ont plus souvent une comorbidité cardiaque, avec une pression artérielle moyenne plus faible et une fréquence cardiaque plus élevée, plus souvent admis pour pathologie cardiaque que les patients non traités par CCD, et moins souvent admis pour pathologie respiratoire, ils ont une PAO2/FIO2 plus faible(ratio inferieur à 200 signe un syndrome de détresse respiratoire aigu), une pression artérielle en CO2 inferieur à la normale et plus faible que le groupe non traité, une creatininémie plus élevée et un score d'apache plus élevé.

Donc globalement, les patients traités par CCD sont dans un état cardiaque et rénal plus sévère et avait donc de base plus de risque de décéder. Il y a donc un biais d'indication et il faut donc absolument prendre en compte cette différence d'état de base en considération lorsque l'on teste si le traitement par CCD diminue le décès à 30 jours.  
Pour cela, nous allons donc réaliser un appariemment sur le score de propension.

# IV - Appariemment sur le score de propension

Le score de propension permet d'avoir pour chaque patient sa probabilité d'être traité par CCD, en fonction de ses charactéristiques de bases. On apparie ensuite un patient traité par CCD avec un patient non traité par CCD qui avait la même probabilité d'être traité que la patient effectivement traité. On se place donc dans la situation d'un essai clinique randomisé ou chaque patient à la même probabilité de recevoir l'un ou l'autre des traitements.

## 1) Sélection des variables à intégrer : tests bivariés

Je dois intégrer dans le score de propension les variables qui sont associées significativement avec le décès à 30 jours ou avec le décès à 30 jours et le traitement par cathétérisme. Pour cela je réalise deux séries de tests bivariés. Une première série testant l'association entre chaque variable et le décès à 30 jours et une deuxième série testant l'association entre chaque variable et le traitement par cathétérisme. Pour rappel, les variables ADLD3P et URIN1 ne sont pas dans l'analyse. J'utilise des modèles de régression logistique à une variable explicative, la variable à expliquer étant soit la variable cathétérisme, soit la variable décès à 30 jours (variables toutes deux binaires) et la variable explicative étant chacune des variables à tester.

Les conditions de validité sont toujours respectées car j'ai 9 classes au maximum (variable CAT1) pour 1918 évènements, donc la condition des 5 à 10 variables par variable est toujours respectée.

2 variables sont liées au décès uniquement : TEMP1 et LIVERHX.

34 variables sont liées au décès et au traitement par cathétérisme : RESP, CARD, NEURO, GASTR, HEMA, SEPS, CAT1, CAT2, CA, DAS2D3PC, DNR1, SURV2MD1, APS1, SCOMA1, WTKILO1, MEANBP1, PAFI1, PACO21, PH1, WBLC1, HEMA1, CREA1, BILI1, ALB1, CARDIOHX, CHFHX, DEMENTHX, PSYCHHX, CHRPULHX, GIBLEDHX, MALIGHX, AGE, INCOME, NINSCLAS.

J'aurai donc 36 variables dans le score de propension.

Les 10 variables liées uniquement au cathétérisme ne sont pas prises en comptes dans le score : RENAL, TRAUMA, RESP1, HRT1, SOD1, IMMUNHX, TRANSHX, AMIHX, SEX, EDU.

## 2) Calcul du score de Propension

Je calcule le score de propension à partir d'un modèle de régression logistique avec comme outcome le traitement par swan ganz et comme variables explicatives toutes les variables liées au décès à 30j uniquement(facteurs pronostiques) ou au décès et au traitement(facteurs de confusion). Je n'intègre pas les variables liées uniquement au traitement pour ne pas perdre de puissance à la fin.

Condition de validité :  
C'est une régression logistique, je dois avoir 5 à 10 évènement par variable. J'ai 36 variables explicatives dans le score de propension dont 5 variables qualitatives : 9 classes pour CAT1, 7 classes pour CAT2, 3 classes pour CA, 4 classes pour INCOME, 6 classes pour NINSCLAS. Ces variables qualitatives seront donc transformées en 9+7+3+4+6-5= 24 variables binaires. Soit l'équivalent de 36-5+24 = 55 variables dans le modèle logistique pour 1918 évènements. J'ai donc plus de 10 évènements par variable explicative, les conditions de validité sont respectées.

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

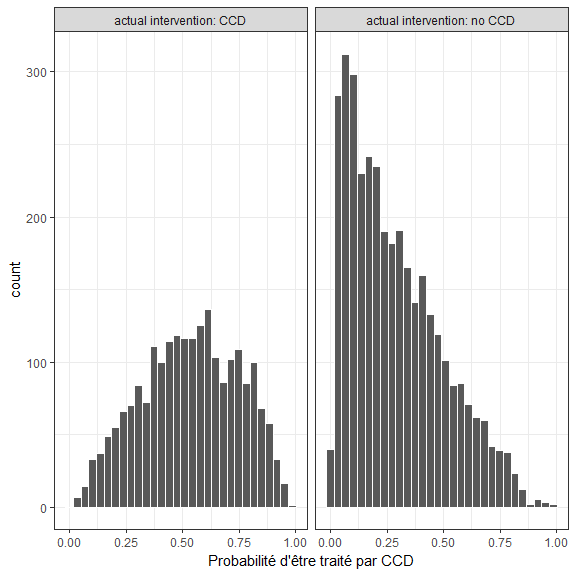


Figure 3. Distribution de la probablité d'être traité par CCD selon le traitement effectif par CCD.

On voit que la distribution du score n'est pas la même parmi les patients ayant été cathétérisé et ceux n'ayant pas été cathétérisé : beaucoup de patient n'ayant pas été cathétérisé avait une faible probabilité d'être cathétérisé (distribution en L), alors que la distribution est plutôt en cloche centré autour de 0.5 pour les patients ayant été cathétérisé (Figure 3). On comprend donc que nécessairement l'appariemment va supprimer de nombreux patients.

## 3) Appariemment sur le score de propension

A partir du score de propension, j'apparie un sujet cathétérisé avec un sujet non cathétérisé ayant un score de propension proche c'est à dire une probabilité d'être cathétérisé proche. Le sujets non appariés sont écartés de l'analyse. Les sujets devrait donc se ressembler dans les 2 groupes.

Deux packages principalement nous permette de réaliser l'appariemment; le package Matching et le package MatchIt. Le package Matching est celui que je trouve le plus simple pour reconstituer le numéro de paire et qui conserve le plus de patients en gardant les memes paramètres.

J'utilise un caliper de 0.2, c'est à dire un seuil d'appariemment de 0.2 x sd(logit(score de propension)). En utilisant le package Matching (fonction Match), sans remise, sans ex aequo, avec un ratio 1:1 et avec un caliper de 0.2, je conserve 3142 patients, 1571 dans chaque groupe.

## 4) Vérification de l'équilibre des variables entre les deux groupes après appariemment

L'appariemment a normalement permis d'avoir des patients globalement comparables en terme de charactéristiques de base, car ayant la même probabilité d'être traité par CCD dans les deux groupes. Il se peut cependant que certaines variables soit mal équilibrée, et c'est ce que nous allons vérifier ici. Seules les variables ayant servie à construire le score de propension doivent être regardées.

Tout d'abord, je transforme mes variables qualitatives en n-k variables binaires, k étant le nombre de classe de la variable qualitative.

J'ai ensuite plusieurs méthodes possibles pour regarder si l'appariemment a équilibré la distribution des variables dans les groupes traité par CCD et non traité par CCD:

* Méthode 1 : je regarde la moyenne de chaque variable en fonction du score de propension (annexe : figure A-2). Je vois donc pour des patients de ces 2 groupes ayant la même probabilité d'être traité, si la moyenne de la variable est semblable. Bien sûr si les courbes se superposent parfaitement, on peut dire que la distribution est semblable dans les deux groupes.Si elles sont disjointes à certaines valeures d'abscisse, nous pouvont dire que pour les individus ayant tel probabilité d'être traité, la moyenne des variables diffèrent. Ainsi dans les courbes présentées en annexe, on voit ainsi que les individus avec une forte probabilité d'être traité ont une moyenne qui semble différer entre les deux groupes pour INCOME, NINCLAS, RESP et SEPS, et que les individus avec une faible probabilité d'être traité ont une moyenne qui semble différer pour les variables RESP, HEMA et DNR1. Or il me semble qu'on ne cherche pas à ce que les individus ayant la même probabilité d'être traité soit exactement semblable 2 à 2 (même si bien sûr c'est idéal) mais plutôt que les populations des groupes traités et non traités soit homogène, comme c'est le cas lorsque l'on randomise. Il me semble donc plus pertinent de voir si globablement les distribution sont les mêmes dans les deux groupes, et les deux méthodes présentées ci dessous répondent à cette question.
* Méthode 2 : je regarde la différence standardisée des moyennes (SMD), c'est à dire la différence entre la moyenne dans le groupe cathétérisé et la moyenne dans le groupe non cathétérisé divisé par la variance commune. J'ai séparé les schémas en variables quantitatives non binaires (figure 4) et variables binaires et qualitatives binarisées (Figure 5) pour plus de lisibilité. En instaurant un seuil de smd à 0.1 comme conseillé dans la littérature, je vois que l'appariemment établi un équilibre entre les deux groupes pour toutes les variables ayant servi à calculer le score de propension.

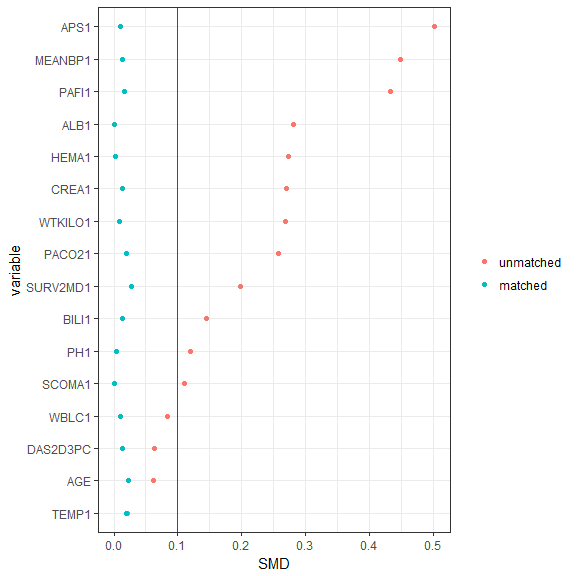


Figure 4. Différence standardisée des moyennes des variables quantitatives.

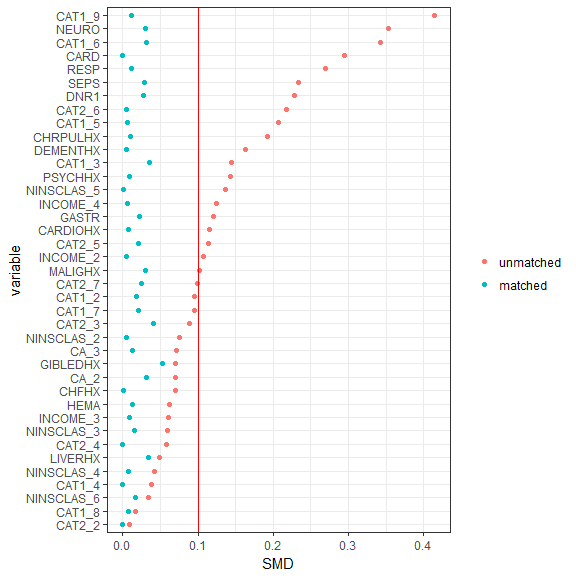


Figure 5. Différence standardisée des moyennes des variables qualitatives (binarisées pour regarder l'équilibre) et des variables binaires.

* Méthode 3 : je regarde les intervalles de confiances des différentes variables, j'ai réalisé les intervalles de confiance des moyennes pour les variables quantitatives (figure 6) et des fréquences de 1 pour les variables binaires et qualitatives binarisées (figure 7). Les variables quantitatives sont standardisées, ce qui permet une lecture plus aisée du graphique des variables quantitatives. Je ne sais pas comment réaliser un équivalent de cette standardisation avec les variables binaires et le graphique est donc moins facilement analysable. A noter que les variables pour lesquelles on n'a pas pu calculer d'intervalle de confiance pour non respect des conditions de validité (np>=5 et n(1-p)>=5) ne sont pas présente dans le graphe. Ce sont un des niveaux de CAT1 et deux des niveaux de CAT2. On voit là aussi que globalement les variables sont plutôt bien équilibrée après appariemment car les intervalles de confiance se recouvrent deux à deux.

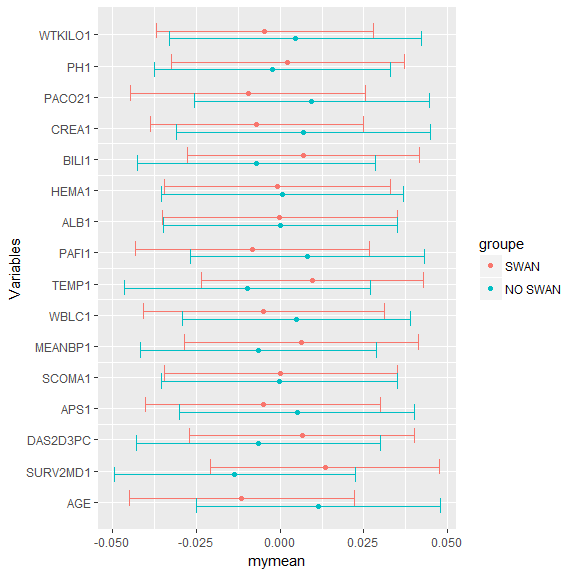


Figure 6. Intervalle de confiance des variables quantitatives utilisées pour calculer le score de propension, en fonction du traitement par CCD.

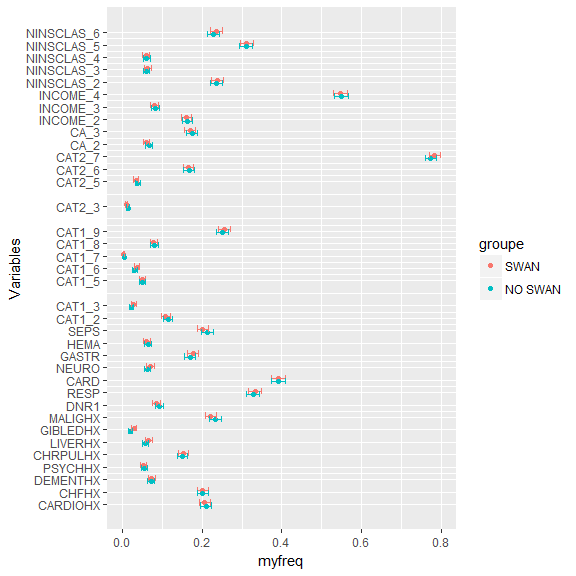


Figure 7. Intervalle de confiance des variables qualitatives et binaires utilisées pour calculer le score de propension, en fonction du traitement par CCD (les variables qualitatives ont été binarisées).

Dans le futur, je préférerai quand même l'analyse de l'équilibre par la méthode smd qui est plus radide à coder et plus facile à lire. C'est dailleurs la méthode que j'ai le plus souvent rencontrée dans les publications.

# V - Analyses dans la population appariée

Pour répondre à la question "le cathétérisme cardiaque droit modifie-t-il la survie à 30 jours?", je peux utiliser deux méthodes différentes. Soit j'utilise une régression logistique conditionnelle, soit j'utilise un modèle de Cox. Dans les deux cas, j'utilise mon échantillon apparié et je stratifie sur le numéro de paire.

## 1) Régression logistique conditionnelle

La variable à expliquer est la mort à 30 jours, la variable explicative est le cathétérisme cardiaque et je stratifie sur le paire.

Je n'ai pas trouvé de source indiquant explicitement quelles sont les conditions de validité de la régression logistique conditionnelle. Par défaut j'ai donc considérée que c'étaient les mêmes conditions que pour la régression logistique et elles sont ici respectées : j'ai plus de 5 à 10 évènements par variable explicative (la stratification sur la paire n'est pas une variable explicative).

Le coefficient du cathétérisme est significatif (pvalue<0.01), le cathétérisme cardiaque droit a donc un effet significatif sur le risaue de décès à 30 jours. le coeffcient vaut 0.17, l'exponentiel du coefficient me permet d'obtenir l'Odds ratio du risque de décès associé au cathétérisme. L'OR est de 1.19 avec un intervalle de confiance à 95% [1.05-1.34]. Le décès n'est pas un évènement rare, je ne peux donc pas interpréter l'OR comme un RR mais je peux dire que le risque de décès est augmenté lorsque le patient est cathétérisé.

## 2) Analyse de survie : Modèle de Cox

Une deuxième manière de réaliser le calcul est par modèle de cox. Dans ce cas il est problématique de prendre en compte l'appariemment par une stratification sur la paire car FE Harrell nous précise dans son livre Regression Modeling Strategies(2nd Ed) p.482 que si le nombre de strates est très grand par rapport au nombre total d'évènement, on perd en efficience(? "Loss of efficiency" dans le texte, je ne sais pas bien comment le traduire). Or on a ici 1569 paires ou strates pour 1055 évènements, ce qui peut être considéré comme un grand nombre de strates relativement au nombre d'évènements bien qu'aucun seuil ne soit précisé. Et de plus FE Harrell nous précise qu'une strate qui ne contient aucun évènement ne contribue pas à l'information et qu'une telle situation doit donc être évitée si possible. Or ici 44% des paires sont sans évènements et ne contribue pas à l'information avec donc j'imagine une perte de puissance. J'utiliserai donc l'option cluster(paire) plutôt que strata(paire) pour prendre en compte l'appariemment dans le modele de cox. Cette option prend en compte le design apparié, et on regarde alors la variance robuste et le test du score robuste. Cette méthode calcule une vraisemblance globale et non pas strate par strate comme lorsque l'on stratifie, évitant donc de perdre de l'information.

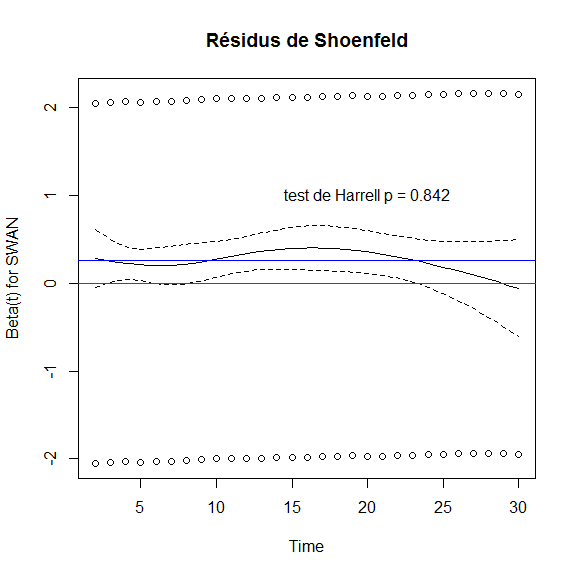


Figure 8. Résidus de Shoenfeld en fonction du temps. P value à partir du test de Harrell.

pvalue   
3.069175e-05

[1] "1.299" "1.297" "1.302"

L'hypothèse des risques proportionnels n'est pas vérifiée. En effet l'intervalle de confiance des résidus de shoenfeld n'incluent pas 0 en tous points (figure 8). Il faut donc en toute rigueur que j'ajoute une variable SWAN dépendante du temps. Cependant je n'ai pas trouvé de transformation qui convienne, aucun coefficient de paramètre dépendant du temps n'étant significatif lorsqu'on l'ajoute au modèle (les résidus de Shoenfeld des modèles ne doivent donc même pas être regardés). Voici les transformations essayeées : log, racine carrée, \*temps, 1/temps, racine cube, carré, cube et puissance de 0.7. J'analyse donc le modèle de Cox sans ajouter de variable dépendante du temps, mais il faudra avoir en tête que le résultat est erroné.

L'ajout d'un cluster sur la paire me permet d'avoir une variance robuste prenant en compte l'appariemment. J'utilise cette variance robuste pour calculer l'intervalle de confiance à 95% du Hazard ratio: 1.236 [1.233-1.239]. Je regarde le test du score robuste pour conclure : p<0.001. Le risque de décès est donc significativement différent selon que le patient est cathétérisé ou non. Et c'est dans le sens d'un risque plus grand chez les patients cathétérisés avec un risque de décès à 30 jours multiplié par 1.236.

Je retrouve cette information graphiquement en traçant une courbe de survie en fonction du traitement par cathéterisme cardiaque droit.

## 3) Représentation graphique de la survie à 30 jours en fonction du traitement par CCD par la méthode de Kaplan-Meier

La méthode de Kaplan Meier permet de représenter graphiquement les courbes de survie (une courbe par groupe).  
Condition de validité de la méthode de Kaplan Meier :  
- censure indépendante de la probabilité de survenue de l'evenement  
- probabilité de survie indépendante du moment de recrutement dans l'étude  
- censure indépendante du groupe

In group CCD = 0

At 10 days, survival[95%CI] 83% [81% - 85%]  
 At 20 days, survival[95%CI] 75% [73% - 77%]  
 At 30 days, survival[95%CI] 70% [68% - 73%]

In group CCD = 1

At 10 days, survival[95%CI] 78% [76% - 80%]  
 At 20 days, survival[95%CI] 69% [66% - 71%]  
 At 30 days, survival[95%CI] 63% [61% - 66%]

Scale for 'colour' is already present. Adding another scale for  
'colour', which will replace the existing scale.

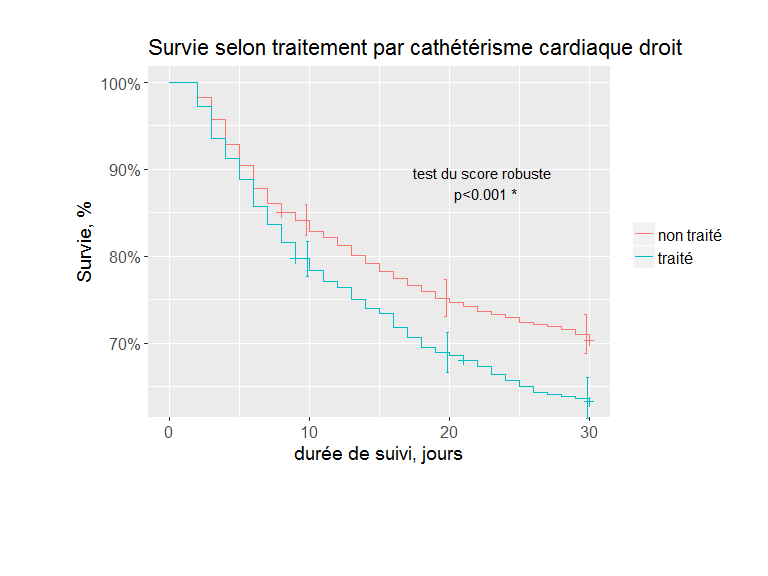


Figure 9. Courbes de survie par la méthode de Kaplan Meier de la survie à 30 jours en fonction du traitement par cathétérisme cardiaque droit. p value calculée par modèle de cox avec cluster sur la paire.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | J0 | J10 | J20 | J30 |
| n.risk | 1571 | 1321 | 1180 | 1114 |
| n.event | 0 | 269 | 129 | 68 |
| surv | 100% | 83 % | 75 % | 70 % |

Table 4. Table de survie des patients non traités par cathétérisme cardiaque droit

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | J0 | J10 | J20 | J30 |
| n.risk | 1571 | 1251 | 1082 | 998 |
| n.event | 0 | 340 | 154 | 83 |
| surv | 100% | 78 % | 69 % | 63 % |

Table 5. Table de survie des patients traités par cathétérisme cardiaque droit

Sur la figure 9, on observe 2 courbes de survie, une pour le groupe non traité pour CCD et une pour le groupe traité par CCD. A chaque temps, les "marches d'escalier" représentent les patients qui ont eu l'évènement, faisant diminuer le nombre de personnes à risque d'évènement(c'est à dire les non répondeurs). Les croix représentent à chaque temps la présence de censure (perdus de vue ou exclus vivant). Par exemple dans le groupe non traité, je passe de près de 85% de survie à J10 à près de 75% de survie à J20 dans le groupe non traité. Dans la table de survie (table 4) je vois que ça correspond à 126 patients ayant eu l'évènement entre J10 et J20. Pour le groupe traité, je lis sur la courbe que je passe de 80% environ de survie à J10 à 70% environ à J20. Dans la table de survie (table 5) cela correspond à 154 patients ayant eu l'évènement. La courbe de Kaplan Meier nous permet d'observer graphiquement que la survie est meilleure lorsque les patients ne sont pas traités par cathétérisme cardiaque droit.

# Conclusion

Nous voulions savoir si le cathétérisme cardiaque droit(CCD) améliorait la survie à 30 jours des patients admis en réanimation. Les données étaients observationnelles, sans randomisation de l'intervention. On ne pouvait donc pas exclure la présence de biais de confusion et d'attrition, qui ont d'ailleurs été mis en évidence dans le devoir : les patients traités par CCD étaient dans un état plus grave. En l'état aucune analyse ne pouvais être faite car elle aurait été biaisée. Afin de tenir compte des charactéristiques de bases nous avons donc réalisé un score de propension et apparié sur ce score. Ainsi chaque individus traité et non traité appariés avaient la même probabilité d'être traités. Après appariemment et en tenant compte de cet appariemment dans l'analyse, nous mettons en évidence une association entre le cathétérisme cardiaque droit par sonde de swan ganz et le risque de mortalité à 30 jours : le cathétérisme cardiaque droit est associé à une augmentation du risque de mortalité. Mais bien que l'utilisation du score de propension augmente le niveau de causalité, on ne peut pas conclure avec le même niveau de causalité qu'un essai randomisé.

# References

Connors, A. F., T. Speroff, N. V. Dawson, C. Thomas, F. E. Harrell, D. Wagner, N. Desbiens, et al. “The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients. SUPPORT Investigators.” JAMA 276, no. 11 (September 18, 1996): 889–97.

Sterne, J. A C, I. R White, J. B Carlin, M. Spratt, P. Royston, M. G Kenward, A. M Wood, and J. R Carpenter. “Multiple Imputation for Missing Data in Epidemiological and Clinical Research: Potential and Pitfalls.” BMJ 338, no. jun29 1 (September 1, 2009): b2393–b2393. <doi:10.1136/bmj.b2393>.

Cours du Dr David Hajage "Evaluation de l’effet d’un traitement en condition réelle d’utilisation (scores de propension et scores pronostiques)", janvier 2017.

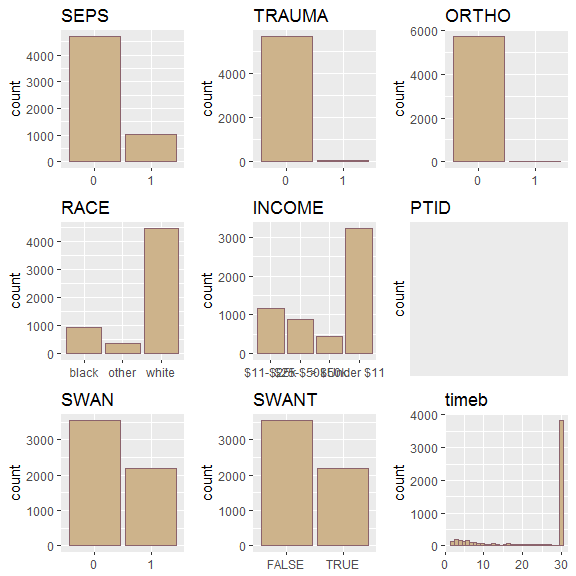
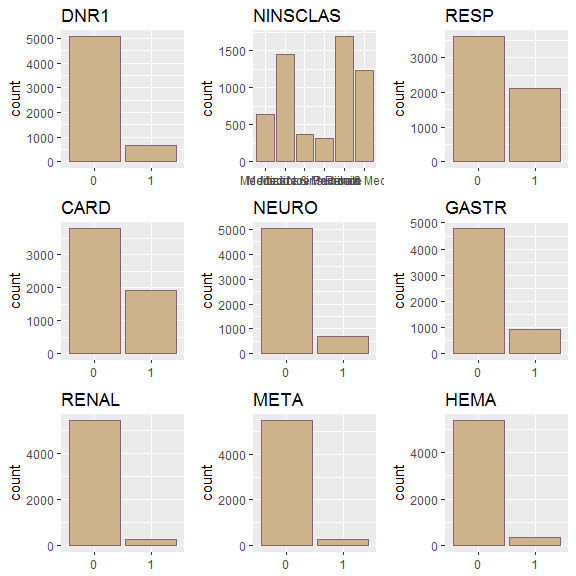
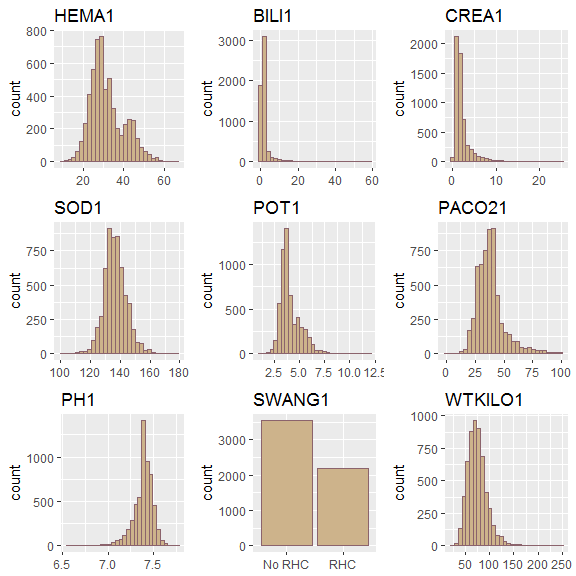
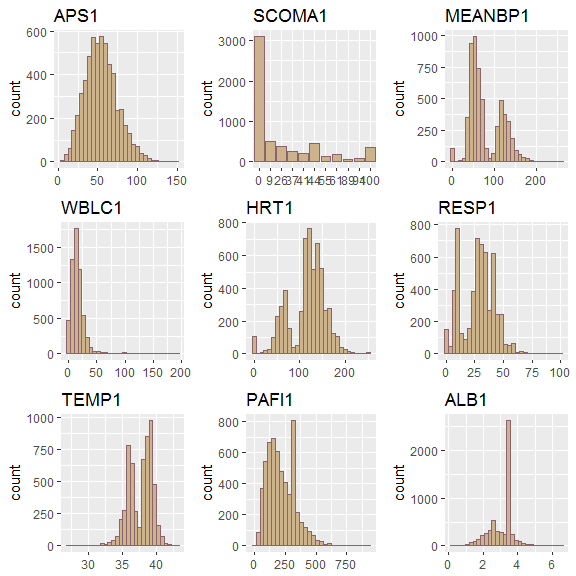
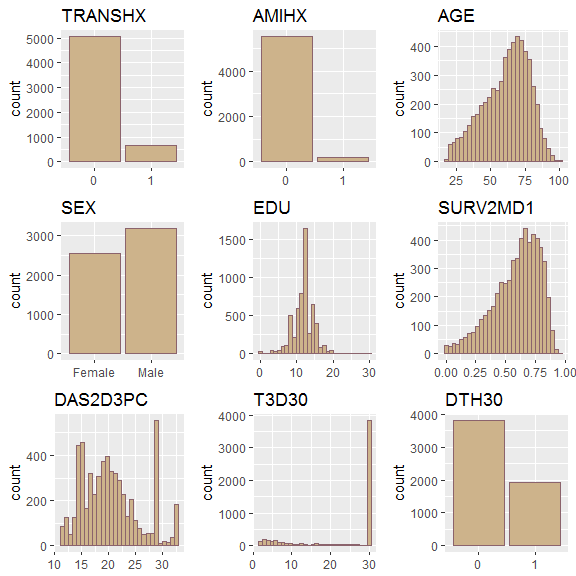
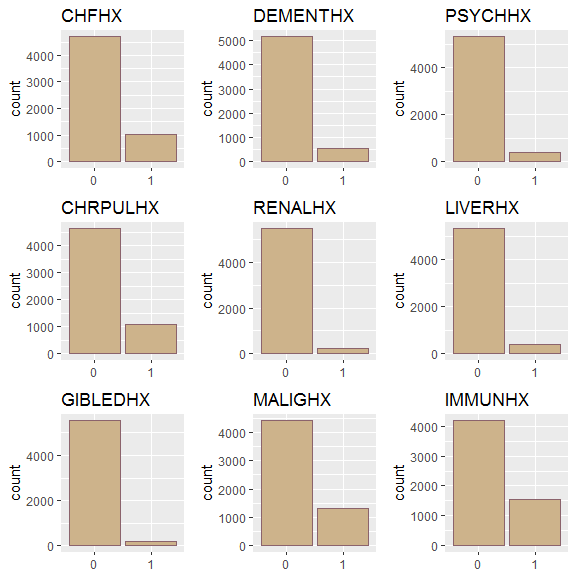
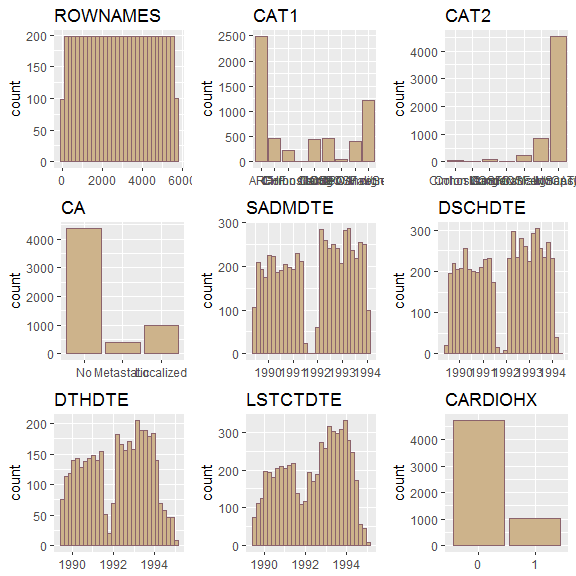
Austin, Peter C. “An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.” Multivariate Behavioral Research 46, no. 3 (May 31, 2011): 399–424. <doi:10.1080/00273171.2011.568786>.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. “MatchIt: Nonparametric Preprocessing for Parametric Causal Inference.” Journal of Statistical Software 42, no. 8 (2011). <doi:10.18637/jss.v042.i08>.

# Annexes

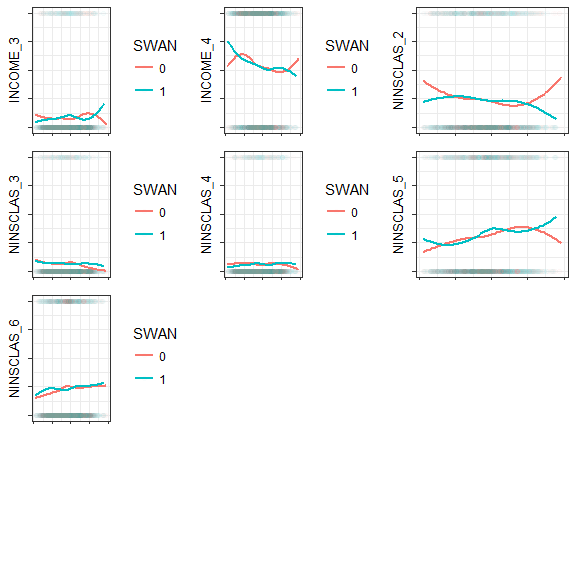
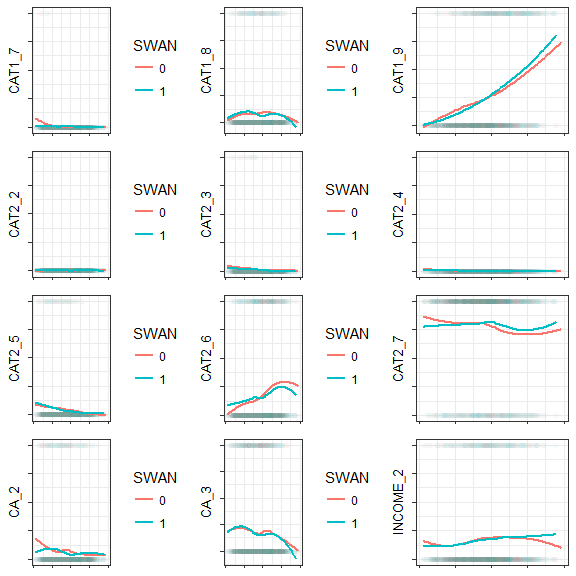
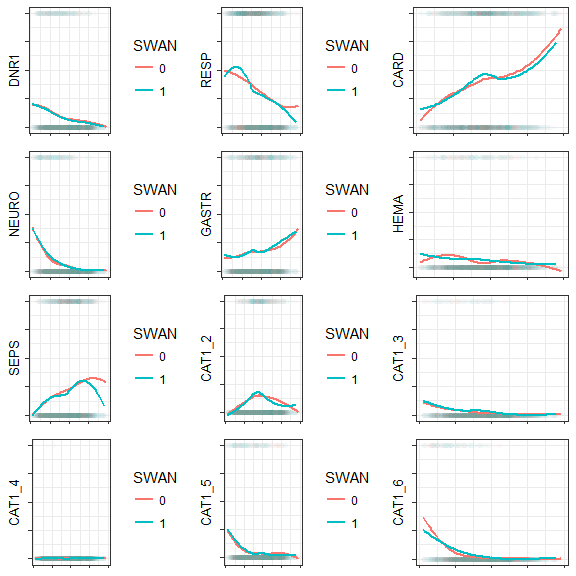
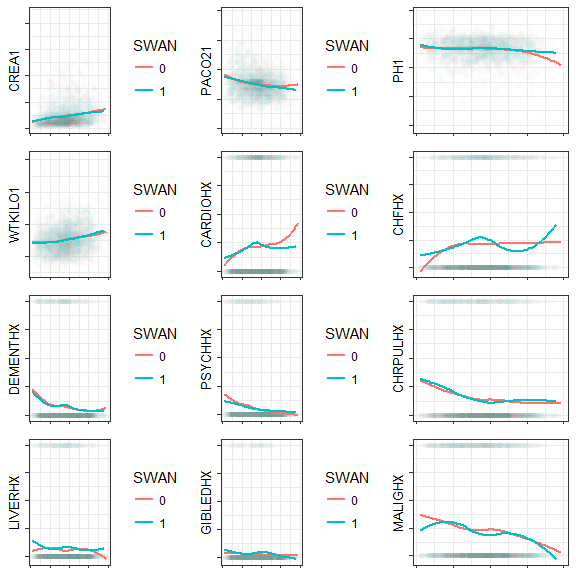
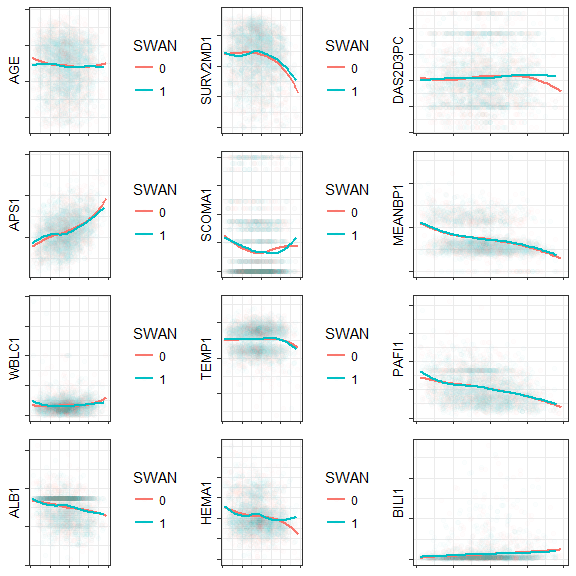
## Figure A-1. Distribution des variables avant modification et imputation

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



[[1]]  
NULL  
  
[[2]]  
NULL  
  
[[3]]  
NULL  
  
[[4]]  
NULL  
  
[[5]]  
NULL  
  
[[6]]  
NULL  
  
[[7]]  
NULL

## Figure A-2. Equilibre des variables après appariemment : moyenne pour chaque variable en fonction du score de propension



[[1]]  
NULL  
  
[[2]]  
NULL  
  
[[3]]  
NULL  
  
[[4]]  
NULL  
  
[[5]]  
NULL