

CONSIGNES POUR L'ENVOI DE VOS DONNÉES AUX BIOSTATISTIENS

Ces consignes sont à respecter pour un traitement rapide et efficace de vos données par les statisticiens de l'Unité de Biostatistiques du CHU de Lille.

1. Le fichier de données

NUMERO	GROUPE	TYPE_EXAM	SEXE	AGE	BMI	TABAC	DIABETE
1	0	1	0	66		0	0
2	0	3	1	62	32	1	0
3	1	3	1	56	25	0	1
4	0	2	1	55	32	0	1
5	1	1	0	55	25	2	0
6	1	2	1	44	26	0	0
7	0	2	0	66	20	2	0
8	1	1	0	49	35	1	0

1.1 La structure

- Fichier anonymisé, juste un numéro de patient.
- En ligne doivent se trouver les individus ; en colonne les variables.
- Pas de colonne vide entre les variables ni de ligne vide entre les patients.

1.2 Le nom des variables

- **Pas d'espaces ni de caractères spéciaux** (é, è, à, ç, ^, <, >, /, -, ?, %, ', etc).
- Seul l'underscore _ (tiret bas de la touche 8 sur un PC) est autorisé pour séparer 2 mots
- <30 caractères
- Ils doivent tous se trouver **sur une seule et même ligne**
- Pas de doublons
- Ne doit pas commencer par un chiffre

1.3 Les données

- Pas de texte (hors noms des variables), tout doit être **codé avec des chiffres**.
 - Variables binaires: 0=Non/absence ; 1=Oui/Présence
 - Variables catégorielle (plusieurs modalités) :
 - Exemple : Type_exam : 1=écho ; 2=scan; 3=IRM; ...
 - Variables où plusieurs réponses sont possibles : faire une variable binaire par possibilité de réponses
 - Exemple : antécédents médicaux : ATCD_CV 0/1, ATCD_respi 0/1, ATCD_HTA 0/1,...
- **Données manquantes = case vide**
- **Evitez de laisser les dates**, calculez l'âge et les délais nécessaires
- Toutes les données doivent être contenues dans **une seule feuille Excel** (sauf cas exceptionnel convenu lors du rendez-vous).

2. Le lexique des variables

Un lexique des variables doit être fourni avec le fichier Excel, soit dans un **fichier Word**, soit dans une **seconde feuille Excel** ou en cas de faible nombre de variable, dans la première ligne du fichier Excel au-dessus des noms de variables.

Pour chaque variable, veuillez indiquer :

- sa signification,
- son unité s'il s'agit d'une variable quantitative,
- la signification des codages pour les variables codées.

Exemple :

Groupe : 0=Témoins 1=Malades

Type_exam : type d'examen réalisé pour le diagnostic de la maladie (1=écho, 2=scan, 3=IRM)

Sexe : 0=Homme 1=Femme

BMI : Indice de masse corporelle kg/m²

Tabac : 0=non-fumeur, 1=fumeur actif, 2= sevré

3. Le plan d'analyse

Le plan d'analyse doit rappeler le titre et contenir une brève description de votre étude ainsi que vos coordonnées (téléphone, mail et spécialité).

Ensuite, il doit reprendre les questions abordées durant le rendez-vous avec les statisticiens. Si vos données sont issues d'un protocole de recherche, il est important de nous fournir la version finale du protocole.

3.1 Le descriptif

Le descriptif de toutes les variables du fichier sur l'ensemble des patients sera réalisé par défaut.

Si vous souhaitez un descriptif de vos sujets par sous-groupe, précisez la variable indiquant ces groupes.

3.2 Les questions

Exemple :

- Comparer les malades aux témoins (GROUPE) selon les paramètres suivants :

- *Age (age)*
- *Sexe (sexe)*
- *Indice de masse corporelle (BMI)*
- *Type d'examen ayant permis le diagnostic (Type_exam)*

Indiquez le nom de la variable tel qu'on le trouvera dans le fichier Excel

Si une variable est à calculer par nos soins pour répondre à l'une de vos questions, indiquez nous la formule à utiliser ou les regroupements de modalité à effectuer.

4. Et ensuite ?

Envoyez-nous vos fichiers à l'adresse suivante : biostatistique@chu-lille.fr.

Pour plus d'informations sur les missions de notre équipe, n'hésitez pas à consulter notre site internet :

<http://seed.chu-lille.fr>

Pour toute demande de rendez-vous pour un nouveau projet, merci de contacter notre secrétariat au 03.20.44.60.38 ou via secretariat-databiostat@chu-lille.fr