

ANALYSE DE DONNÉES DISCRÈTES
APPLIQUÉE À L'EFFICACITÉ
DE LA SURVEILLANCE DE LA CONDITION FOETALE

Par
Caroline Anctil, B.Sc.
Mai 1996

Thèse présentée à l'Ecole des Etudes Supérieures
de l'Université d'Ottawa pour l'obtention
de la Maîtrise ès Sciences en Mathématiques¹

© Copyright 1996
par Caroline Anctil, B.Sc., Ottawa, Canada

¹Le programme de la maîtrise est conjoint avec l'Université Carleton et est administré par l'Institut de Mathématiques et de Statistiques d'Ottawa-Carleton



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-15587-0

Canada



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

Résumé

Dans le domaine de la recherche, les observations sont souvent de type qualitatif. Des variables discrètes sont définies pour représenter ces données. Afin d'analyser ce type de variables, différentes techniques statistiques sont disponibles.

Le but de cette thèse est d'expliquer, de développer et d'utiliser des tests statistiques basés sur différentes méthodes d'analyse appliquées à une expérience médicale lorsque les données sont de type qualitatif. Par la suite, ces techniques seront comparées entre elles. De même, tout au long de ce document, des remarques seront apportées sur les problèmes encourus lors du développement et de l'application de ces statistiques.

Parmi les techniques d'analyse choisies, la méthode du maximum de vraisemblance sera employée pour établir une statistique de test. De plus, le test de McNemar, la régression logistique et le ré-échantillonnage seront utilisés. Des tests seront développés pour vérifier des hypothèses simples et multiples. Finalement, les techniques statistiques seront évaluées dans le cas d'une analyse univariée de même que dans le cas d'une analyse multivariée, lorsque la situation le permettra.

Remerciements

Je tiens à remercier très chaleureusement mon directeur de thèse, André Dabrowski, pour avoir dirigé ma recherche et m'avoir prodigué des conseils utiles au succès de ma maîtrise. J'ai grandement apprécié sa patience et son intérêt soutenu pour mon travail.

De même, je voudrais remercier Dr. Carl Nimrod, obstétricien-gynécologue chef à l'Hôpital Général d'Ottawa, et son équipe du département de périnatalogie pour leur aide précieuse durant le développement de l'expérience ayant permise d'acquérir les données nécessaires pour l'analyse statistique. Je voudrais également souligner la grande coopération de Dr. Lawrence Oppenheimer et de Dr. Nigel Simpson qui a rendue possible le succès de ce projet.

Je voudrais également remercier l'Université d'Ottawa pour le support financier accordé tout au long du programme.

Je tiens finalement à souligner le support constant de mon conjoint et de mes amis tout au long de mes études universitaires. Je leur en suis très reconnaissante.

Dédicace

A deux personnes très importantes pour moi,

Jonathan et Pierre

Table des matières

Résumé	ii
Remerciements	iii
Dédicace	iv
1 Introduction	1
2 Description de l'expérience	3
3 Plan de l'expérience	7
4 Hypothèses à vérifier	10
5 Techniques d'analyse utilisées	14
5.1 Test de Student (test t)	16
5.1.1 Remarques	17
5.2 Méthode du maximum de vraisemblance	18
5.2.1 Comparaisons simples	19
5.2.2 Remarques	22
5.2.3 Comparaisons multiples	23
5.3 Test de McNemar	25
5.4 Régression logistique	27
5.4.1 Modèle	27
5.4.2 Comparaisons simples	30

5.4.3	Comparaisons multiples	34
5.4.4	Estimation	38
5.4.5	Remarques	39
5.5	Ré-échantillonnage	41
6	Analyse des résultats et comparaisons des méthodes statistiques	46
6.1	Comparaisons simples	46
6.2	Comparaisons multiples	49
6.3	Remarques	50
6.4	Conclusions	50
A	Description technique du matériel employé	52
B	Présentation des tracés	54
C	Exemplaire du questionnaire utilisé pour l'expérience	58
D	Liste des données originales	60
E	Programmes informatiques	67
F	Table de contingence	84
G	Estimateurs du maximum de vraisemblance calculés avec MATHE-	
	MATICA	85
H	Intervalles de confiance de Clopper-Pearson	87
I	Présentation des résultats	89
I.1	Analyse univariée (Comparaisons simples)	90
I.2	Analyse univariée (Comparaisons multiples)	94
I.3	Résultats du test de McNemar	97
	Bibliographie	98

Chapitre 1

Introduction

Lors d'une analyse statistique, il faut toujours tenir compte du type de données à analyser. Les variables peuvent être soit quantitatives, soit qualitatives. Les méthodes d'analyse à utiliser sont déterminées par le type des variables de l'étude.

Ce document a pour but de présenter différentes méthodes statistiques applicables choisies pour l'analyse de variables discrètes de type qualitatif. Ces techniques seront appliquées aux données venant d'une expérience réalisée au département de périnatalogie de l'Hôpital Général d'Ottawa concernant l'efficacité de techniques de surveillance de la condition foetale lors d'un accouchement. Le but de cette expérience était d'ordonner 3 techniques de contrôle selon leur rendement.

Tout d'abord, les hypothèses seront testées à l'aide d'un test apparié de Student (test t). Par la suite, la technique du maximum de vraisemblance qui vise à créer une statistique de test en utilisant des estimateurs pour les paramètres du modèle étudié qui sont "les plus vraisemblables" d'avoir produit les données observées sera appliquée. Également, une brève explication du test de McNemar sera apportée. Ensuite, la régression logistique sera utilisée afin de vérifier s'il existe une relation entre les variables (indépendantes et dépendantes) du modèle. Finalement, le ré-échantillonnage sera mis en pratique afin de conclure si l'échantillon prélevé est caractéristique de la

population cible par rapport à un grand nombre d'échantillons possibles qui auraient pu être tirés.

L'évaluation des outils statistiques précédemment cités (à l'exception du test de Student), c'est-à-dire, le rapport du maximum de vraisemblance, le test de McNemar, la régression logistique et le ré-échantillonnage sera faite lorsque les 3 techniques de contrôle seront d'abord comparées 2 à la fois. Ces comparaisons sont appelées des comparaisons simples. Une seconde évaluation sera également entreprise lorsque ces mêmes techniques seront comparées simultanément (les 3 à la fois), c'est-à-dire qu'une comparaison multiple sera vérifiée. Cette évaluation comprendra l'analyse des résultats obtenus en les comparant entre les différentes techniques statistiques employées.

Finalement, une analyse multivariée viendra clore l'évaluation des techniques statistiques pour les cas où il est possible de bâtir un test.

Le prochain chapitre présente l'expérience qui a permis de récolter les données qui ont servi à la présente analyse. Cette expérience s'est déroulée au département d'obstétrique et de gynécologie de l'Hôpital Général d'Ottawa.

Chapitre 2

Description de l'expérience

Le but de l'expérience était de comparer différentes procédures d'évaluation de la condition foetale lors d'un accouchement. Trois techniques ont été utilisées: A) le diagnostic (calcul des battements cardiaques foetaux) avec un stéthoscope appuyé sur le ventre de la mère, M) le diagnostic avec un stéthoscope aidé d'un compteur électronique ("meter") des battements du coeur foetal et V) le diagnostic avec une représentation visuelle des courbes des battements du coeur foetal et d'une contraction.¹

Afin d'évaluer ces 3 procédés, 6 tracés caractéristiques représentant les battements cardiaques foetaux ont été extraites d'une base de données électronique de tracés archivés à l'unité de gynécologie de l'Hôpital Général d'Ottawa. Chacun de ces exemples présente 2 courbes d'une durée de 3 minutes. La première de ces courbes représente le taux des battements cardiaques foetaux mesuré en battements par minute (bpm). La deuxième courbe (appelée TOCO) symbolise la pression exercée par l'utérus et le pic caractérise une contraction. (Voir l'Appendice B). A l'aide d'un programme informatique développé spécialement pour l'expérience (Voir l'appendice A pour une description du programme), ces 6 tracés ont été présentés

¹L'ordinateur utilisé pour présenter ces techniques était équipé d'une carte graphique et de haut-parleurs afin de reproduire le son et l'aspect graphique des données. Voir l'Appendice A pour plus de détails sur le logiciel utilisé appelé TRACER.

selon les 3 modes différents correspondant aux 3 techniques précédemment mentionnées. Donc, 18 versions des courbes ont ainsi été obtenues. Ces 18 versions ont alors été présentées dans un ordre aléatoire différent à chacun des 30 résidents et infirmières et en gynécologie de l'Hôpital qui ont fait partie de l'expérience. La taille d'échantillon a été fixée à 30 personnes car ce nombre représentait le personnel qualifié du département disponible pour participer à l'épreuve.

Lors de l'expérience, ces personnes devaient remplir un questionnaire pour chacune des 18 représentations sur les paramètres suivants,

Q1. le taux de base des battements cardiaques foetaux par minute (bpm)

Q2. la variabilité de base des battements cardiaques foetaux

2a. réduite (< 6 bpm)

2b. normale (6-25 bpm)

2c. excessive (> 25 bpm)

Q3. les changements périodiques des battements cardiaques foetaux

3a. pas de changement présent

3b. accélération des battements cardiaques

3c. décélération des battements cardiaques
(avec récupération avant la fin de la contraction)

3d. décélération des battements cardiaques
(avec récupération après la fin de la contraction)

Q4. l'interprétation de la courbe

4a. rassurante

4b. non-rassurante

Le questionnaire était présenté sous forme de grille (Voir l'Appendice C) où chaque participant devait cocher la case qu'il jugeait correspondre à la bonne réponse. Chaque item précédent pour chaque question correspond à une case du questionnaire et sont des réponses possibles à la question posée. Donc, les données observées sont catégoriques ("categorical data").

La question Q1 demande au participant d'enregistrer la fréquence de base des battements cardiaques par minute du fœtus selon la courbe présentée. Cette valeur peut être dérivée à l'aide de la première courbe des tracés. La question Q2 réfère aux fluctuations des battements du cœur fœtal. Si le taux est très variable, le fœtus peut être en détresse et il faut alors immédiatement prendre des procédures pour mettre rapidement le bébé au monde. Ces situations surviennent parfois lors de fortes contractions de l'utérus. La question Q3 consiste à analyser le comportement du fœtus face aux contractions, c'est-à-dire si celles-ci perturbent ou non ses battements cardiaques. Les contractions peuvent décélérer le taux des battements cardiaques. Dans ce cas, il faut vérifier si le fœtus semble récupérer avant ou après la fin de la contraction. La question Q4 demande une interprétation globale de la situation pour la période présentée (incluant une contraction) à savoir, si la situation semble rassurante ou non en ce qui concerne la condition foetale.

Les études statistiques qui seront présentées dans ce document porteront exclusivement sur les questions Q2, Q3 et Q4. Ces questions représentent chacune une variable discrète et qualitative. (Le taux de base des battements cardiaques foetaux est une variable continue.) A partir de celles-ci, il est possible de déterminer si les réponses sont correctes ou non. Ces bonnes réponses sont fournies par les médecins spécialistes du département de gynécologie de l'Hôpital. Donc, les variables dépendantes de l'étude sont dénotées par

B pour la variabilité de base des battements foetaux

P pour les changements périodiques

I pour l'interprétation

et prennent les valeurs suivantes,

$$B = \begin{cases} 1 & \text{si réponse à la question 2 est correcte} \\ 0 & \text{si réponse à la question 2 est incorrecte} \end{cases}$$

$$P = \begin{cases} 1 & \text{si réponse à la question 3 est correcte} \\ 0 & \text{si réponse à la question 3 est incorrecte} \end{cases}$$

$$I = \begin{cases} 1 & \text{si réponse à la question 4 est correcte} \\ 0 & \text{si réponse à la question 4 est incorrecte} \end{cases}$$

(Voir l'Appendice D pour un aperçu des données originales).

Chaque catégorie définissant les questions a été codifiée. Ces codes sont identifiés dans l'exemplaire du questionnaire qui se trouve en Appendice C. Avec ces codes, il est simple de convertir les données qualitatives en données numériques binaires (0 et 1) à l'aide d'un programme informatique.

Se basant sur l'information précédemment fournie, la section qui suit décrira le plan de l'expérience.

Chapitre 3

Plan de l'expérience

Pour l'expérience exposée précédemment, chacun des 30 individus y participant devait essayer de donner la bonne réponse à 3 questions posées (variables B , P et I) appliquée à 3 méthodes de contrôle (traitements A , M et V). Donc, l'expérience a été développée selon un plan de mesures répétées car, pour chaque sujet, 3 "traitements" ont été administrés. Le modèle de cette expérience est donc donné par

$$\bar{Y} = \bar{\mu} + \bar{\varepsilon} \quad (1)$$

ou

$$\begin{bmatrix} A \\ M \\ V \end{bmatrix} = \begin{bmatrix} p(A) \\ p(M) \\ p(V) \end{bmatrix} + \begin{bmatrix} \varepsilon_A \\ \varepsilon_M \\ \varepsilon_V \end{bmatrix} \quad (2)$$

Les paramètres $p(A)$, $p(M)$ et $p(V)$ sont les proportions de succès associées avec chacune des méthodes de contrôle à tester. La proportion de succès d'une question est calculée comme étant le nombre de personnes ayant une bonne réponse (la fréquence des 1) à la question posée divisé par le nombre total de personnes participant à

l'expérience. Pour l'expérience décrite, $n = 30$ pour tous les cas.

La proportion de succès pour la méthode utilisant seulement le stéthoscope est dénotée par $p(A)$ (A pour "audio"), pour la méthode utilisant un stéthoscope aidé d'un compteur électronique, elle est appelée $p(M)$ (M pour "meter") et pour la méthode utilisant une représentation graphique visuelle, elle est $p(V)$.

Ces proportions sont obtenues pour chacune des variables dépendantes observées. Les observations sont B_1, \dots, B_{30} , P_1, \dots, P_{30} et I_1, \dots, I_{30} pour les variables B , P et I , respectivement. Donc, pour la variable B , la proportion $p(A)$ devient $p(A_B)$, pour la variable P , la proportion $p(A)$ devient $p(A_P)$ et pour la variable I , la proportion $p(A)$ devient $p(A_I)$. Les proportions $p(M)$ et $p(V)$ sont décrites similairement selon ce schéma. Donc, la définition des proportions de succès à partir des observations sont

$$p(A_B) = \Pr(B_i = 1)$$

$$p(A_P) = \Pr(P_i = 1)$$

$$p(A_I) = \Pr(I_i = 1)$$

lorsque toutes les variables sont évaluées selon la technique A . Similairement, pour la technique M ,

$$p(M_B) = \Pr(B_i = 1)$$

$$p(M_P) = \Pr(P_i = 1)$$

$$p(M_I) = \Pr(I_i = 1)$$

et finalement, pour la technique V ,

$$p(V_B) = \Pr(B_i = 1)$$

$$p(V_P) = \Pr(P_i = 1)$$

$$p(V_I) = \Pr(I_i = 1)$$

Toutes ces proportions doivent être calculées pour les 6 tracés de l'étude car les hypothèses seront basées sur ces proportions de succès. D'ailleurs, le chapitre suivant détaillera ces hypothèses qui seront vérifiées pour cette étude.

Chapitre 4

Hypothèses à vérifier

Pour l'expérience décrite, la question d'intérêt est de vérifier laquelle des 3 méthodes d'évaluation précédemment mentionnées (A , M et V) semble la meilleure. La réponse à cette question peut être obtenue en testant 2 différents types d'hypothèses; les comparaisons simples et les comparaisons multiples. Ces comparaisons seront faites sur les proportions de succès de chacune des questions. Pour ordonner les méthodes de contrôle selon leur efficacité, une hypothèse alternative unilatérale ("one-sided") est opposée à l'hypothèse nulle basée sur l'égalité des proportions de succès.

Les comparaisons simples consistent à comparer les proportions de succès des méthodes d'évaluation en les prenant 2 à la fois. C'est-à-dire que la probabilité que l'intervalle de confiance ne contienne pas la moyenne, pour chaque comparaison, est de $\alpha = 0.05$. Donc, lorsqu'il y a plus d'une comparaison simple à vérifier pour une expérience donnée, l'erreur de type I est alors grandement augmentée (proportionnellement au nombre de comparaisons simples). C'est pour cette raison qu'il est parfois avantageux de considérer des hypothèses multiples. Pour l'expérience décrite précédemment, il y a un total de 3 paramètres à tester. Donc, les comparaisons multiples consistent à comparer les proportions de succès des 3 méthodes d'évaluation simultanément (les 3 à la fois). C'est-à-dire que la probabilité que les intervalles de confiance ne contiennent pas toutes les moyennes est $\alpha = 0.05$.

Les tableaux qui suivent montrent la série d'hypothèses à tester. Le premier présente les hypothèses simples (lorsque les méthodes de contrôle sont comparées 2 à la fois) tandis que le deuxième présente les comparaisons multiples (lorsque les méthodes de contrôle sont comparées 3 à la fois.). Ces comparaisons seront faites pour les 6 tracés de l'étude.

	H_0	H_1
Variable B	$p(A_B) = p(M_B)$	$p(A_B) < p(M_B)$
	$p(M_B) = p(V_B)$	$p(M_B) < p(V_B)$
Variable P	$p(A_P) = p(M_P)$	$p(A_P) < p(M_P)$
	$p(M_P) = p(V_P)$	$p(M_P) < p(V_P)$
Variable I	$p(A_I) = p(M_I)$	$p(A_I) < p(M_I)$
	$p(M_I) = p(V_I)$	$p(M_I) < p(V_I)$

Tableau 1: Ce tableau présente les hypothèses à tester pour les 3 variables de l'étude B , P et I . Ces comparaisons simples seront faites pour les 6 tracés ayant été présentés pendant l'expérience.

	H_0	H_1
Variable B	$p(A_B) = p(M_B) = p(V_B)$	$p(A_B) < p(M_B)$ ou $p(M_B) < p(V_B)$
Variable P	$p(A_P) = p(M_P) = p(V_P)$	$p(A_P) < p(M_P)$ ou $p(M_P) < p(V_P)$
Variable I	$p(A_I) = p(M_I) = p(V_I)$	$p(A_I) < p(M_I)$ ou $p(M_I) < p(V_I)$

Tableau 2: Ce tableau présente les hypothèses à tester pour les 3 variables de l'étude B , P et I . Ces comparaisons multiples seront faites pour les 6 tracés ayant été présentés pendant l'expérience.

Lorsqu'une analyse multivariée sera considérée, les variables B , P et I seront alors pris comme un vecteur de variables aléatoires dépendantes. Donc, les proportions de succès sont alors définies comme étant un vecteur de proportions tel que

$$p(A) = \begin{bmatrix} p(A_B) \\ p(A_P) \\ p(A_I) \end{bmatrix} \quad (3)$$

$$p(M) = \begin{bmatrix} p(M_B) \\ p(M_P) \\ p(M_I) \end{bmatrix} \quad (4)$$

$$p(V) = \begin{bmatrix} p(V_B) \\ p(V_P) \\ p(V_I) \end{bmatrix} \quad (5)$$

Les hypothèses à vérifier sont résumées dans les 2 tableaux qui suivent.

H_0	H_1
$p(A) = p(M)$	$p(A) < p(M)$
$p(M) = p(V)$	$p(M) < p(V)$

Tableau 3: Ce tableau présente les hypothèses à tester pour le cas d'une analyse multivariée regroupant les 3 variables dépendantes B , P et I . Ces comparaisons simples seront faites pour les 6 tracés ayant été présentés pendant l'expérience.

H_0	H_1
$p(A) = p(M) = p(V)$	$p(A) < p(M)$ ou $p(M) < p(V)$

Tableau 4: Ce tableau présente l'hypothèse à tester pour le cas d'une analyse multivariée regroupant les 3 variables dépendantes B , P et I . Cette comparaison multiple sera faite pour les 6 tracés ayant été présentés pendant l'expérience.

Les prochaines sections présenteront les différentes techniques statistiques choisies pour l'analyse de données discrètes. Avant tout, les hypothèses seront testées en utilisant un test de Student. Par la suite, la méthode du maximum de vraisemblance sera expliquée et une statistique de test basée sur cette méthode sera développée. Suivra, une brève description du test de McNemar. Finalement, la régression logistique sera exposée de même que le procédé de ré-échantillonnage.

Chapitre 5

Techniques d'analyse utilisées

Les techniques statistiques qui seront présentées dans ce chapitre visent à analyser les données recueillies lors de l'expérience visant à comparer 3 techniques de contrôle de la surveillance foetale. Premièrement, le test de Student, communément appelé test t , sera employé pour vérifier les hypothèses simples données dans le chapitre précédent. Ce test est ici effectué étant donné qu'il est très largement utilisé dans le domaine scientifique et que parfois il l'est à tort. La situation présente en est un exemple. Donc, son applicabilité à la présente étude sera ici critiquée. Vu que le test de Student est utilisé uniquement pour vérifier des hypothèses simples, avec la méthode du maximum de vraisemblance, une tentative sera faite afin de développer un test pour vérifier les hypothèses multiples. Il semble que selon cette méthode, cette tâche en soit une de taille.

Egalement, une brève description du test de McNemar sera apportée car des résultats de ce test sont utilisés dans l'analyse.

Etant donné ces insuccès, des statistiques seront établies pour tester les hypothèses simples et multiples pour une analyse univariée à l'aide de la régression logistique, spécialisée dans l'analyse de variables dépendantes discrètes. De plus, des remarques seront faites sur la possibilité de développer un test pour le cas d'une analyse multivariée.

Finalement, une dernière alternative disponible concerne la technique du ré-échantillonnage. Il semble que cette méthode soit relativement simple à appliquer. Par contre, une attention particulière doit être apportée dans le cas présent de mesures répétées.

Donc, les 5 prochaines sections de ce chapitre présenteront en détails le développement des tests selon chacune des techniques statistiques ci-haut mentionnées de même que les problèmes rencontrés.

5.1 Test de Student (test t)

Supposons un échantillon aléatoire de taille n sur lequel 2 mesures sont prises sur 2 "traitements" différents; $X_{11}, X_{12}, \dots, X_{1n}$ et $X_{21}, X_{22}, \dots, X_{2n}$. Le but est de vérifier s'il y a une différence entre les 2 traitements. Donc, l'hypothèse nulle

$$H_0 : \text{moyenne}_1 = \text{moyenne}_2$$

doit être confrontée à l'hypothèse alternative composée

$$H_1 : \text{moyenne}_1 < \text{moyenne}_2$$

En général, le test de Student est utilisé pour tester ce type d'hypothèse.

Le but de la présente expérience est de vérifier si les 3 méthodes de surveillance foetale sont équivalentes. Elles seront comparées 2 à la fois en utilisant le test de Student. Le test prendra la forme suivante,

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \quad (6)$$

où \bar{d} est la moyenne des différences entre les 2 ensembles de n observations appariées et s_d est l'écart-type de ces différences.

L'hypothèse H_0 sera rejetée seulement si

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \leq t_{\alpha}(n-1) \quad (7)$$

Ce test est effectué à un niveau de signification α et $t_\alpha(n - 1)$ est un point critique de la densité de Student à $n - 1$ degrés de liberté. Finalement, la *p-value* est donnée par $\Pr(t \leq t_\alpha(n - 1) | H_0)$. Les résultats de ce test se trouve à l'Appendice I.

5.1.1 Remarques

Ce test est très souvent utilisé lorsque l'on veut comparer si 2 moyennes sont égales ou non. Par contre, pour appliquer ce test, il faut d'abord que les données soient continues. La symétrie et la normalité des observations sont également souhaitables. Pour la présente expérience, les données sont très concentrées. Les différences obtenues pour calculer la statistique \bar{d} prennent les valeurs -1, 0 ou 1 étant donné que les observations initiales sont de types catégoriques ayant les valeurs 0 et 1. Il est donc peu probable d'obtenir les conditions nécessaires pour l'applicabilité de ce test avec une forte concentration de données de ce genre. Donc, les conclusions tirées d'un tel test appliqué à ce type de données ne sont pas valables.

C'est pour cette raison que la prochaine section présentera une méthode statistique plus adaptée pour ce type de données; la méthode du maximum de vraisemblance.

5.2 Méthode du maximum de vraisemblance

La méthode du maximum de vraisemblance consiste à trouver un estimateur pour chaque paramètre du modèle étudié qui est “le plus probable (vraisemblable)” d’avoir produit les données observées ([1], chapitre 3).

Supposons un échantillon aléatoire X_1, X_2, \dots, X_n ayant une densité de probabilité connue $f(x; \theta_1, \theta_2, \dots, \theta_m)$ qui dépend de paramètres inconnus $\theta_1, \theta_2, \dots, \theta_m$. Le paramètre $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ doit être restreint à un espace paramétrique, Θ . La densité de probabilité conjointe de X_1, X_2, \dots, X_n , prise comme une fonction de θ , est appelée fonction de vraisemblance et est donnée comme suit

$$L(\theta) = L(\theta_1, \theta_2, \dots, \theta_m) = \prod_{i=1}^n f(x_i; \theta_1, \dots, \theta_m) = \prod_{i=1}^n f(x_i; \theta) \quad (8)$$

avec $\theta \in \Theta$.

L’estimateur du maximum de vraisemblance pour chaque θ_i ($i = 1, \dots, m$) est $\hat{\theta}_i = g_i(X_1, X_2, \dots, X_n)$. Ces estimateurs maximisent la fonction de vraisemblance $L(\theta)$. De même, les estimations du maximum de vraisemblance, $g_i(x_1, x_2, \dots, x_n)$, sont obtenues à partir des données de l’échantillon. Les estimateurs sont calculés en prenant d’abord le logarithme naturel de $L(\theta)$ et ensuite, la dérivée première selon chaque paramètre est effectuée et égalée à zéro afin de trouver une solution en fonction des x_i pour chaque θ_i .

Afin de tester une hypothèse $H_0 : \theta \in \Theta$ opposée à $H_1 : \theta \notin \Theta$, il est également nécessaire de trouver les estimateurs du maximum de vraisemblance sous H_0 , dénotés par $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_m)$. Pour ce faire, il faut définir la fonction de vraisemblance restreinte sous H_0 . De cette façon, il est possible d’avoir une statistique pour tester l’hypothèse nulle en construisant le rapport de vraisemblance $T = L(\hat{\theta})/L_{H_0}(\hat{\theta})$.

Donc, l'hypothèse H_0 sera rejetée si T est petit.

Cette technique peut maintenant être appliquée à l'expérience précédemment décrite.

5.2.1 Comparaisons simples

Pour chacune des questions (Q2, Q3 et Q4) de type qualitatif discret, il est possible de définir une réponse comme étant "correcte (c)" ou "incorrecte (i)". Pour une paire de questions, pour un même tracé, les données recueillies peuvent être résumées à l'aide d'une table de contingence de dimension 2×2 (Voir la Table 8 en Appendice F).

Donc, pour chaque personne participant à l'épreuve, 4 résultats différents sont possibles;

Cas 1: $\Pr_v(\text{Méthode 1} = \text{correcte}, \text{Méthode 2} = \text{correcte}) = p_{cc}$

Cas 2: $\Pr_v(\text{Méthode 1} = \text{correcte}, \text{Méthode 2} = \text{incorrecte}) = p_{ci}$

Cas 3: $\Pr_v(\text{Méthode 1} = \text{incorrecte}, \text{Méthode 2} = \text{correcte}) = p_{ic}$

Cas 4: $\Pr_v(\text{Méthode 1} = \text{incorrecte}, \text{Méthode 2} = \text{incorrecte}) = p_{ii}$

où v est la variable mesurée et est soit B , P ou I et les méthodes sont A , M ou V comparées 2 à la fois. Chacune des probabilités est définie à partir du nombre de personnes ayant un résultat $x \in \{c, i\}$ avec la première technique et un résultat $y \in \{c, i\}$ avec la deuxième technique. Le symbole n_{xy} représente ce nombre. La probabilité associée est $p_{xy} = n_{xy}/n$ où $n = 30$. La somme des 4 proportions ainsi calculée est 1. Le but est de vérifier si les proportions de succès de 2 méthodes sont les mêmes. Donc, l'hypothèse nulle est $H_0 : p_c = p_{.c}$. Par définition, $p_c = p_{cc} + p_{ci}$ et $p_{.c} = p_{cc} + p_{ic}$. Par exemple, H_0 prend la forme $H_0 : p(A) = p(M)$ lorsque les méthodes A et M sont confrontées (Voir la Section 4 pour la liste exhaustive des

hypothèses à vérifier).

Donc, la densité de probabilité conjointe reliée à cette expérience

$$\Pr(X_{cc} = n_{cc}, X_{ci} = n_{ci}, X_{ic} = n_{ic}, X_{ii} = n_{ii})$$

peut donc être définie par la distribution multinomiale qui est donnée par,

$$\frac{n!}{n_{cc}!n_{ci}!n_{ic}!n_{ii}!} p_{cc}^{n_{cc}} p_{ci}^{n_{ci}} p_{ic}^{n_{ic}} p_{ii}^{n_{ii}} \quad (9)$$

avec $\sum n_{xy} = n$ et $\Theta = \mathbf{p} = (p_{cc}, p_{ci}, p_{ic}, p_{ii})$ où chaque probabilité se situe entre 0 et 1 (i.e. $\{0 \leq p_{xy} \leq 1\}$). L'équation 9 représente la fonction de vraisemblance et est dénotée par $L(\mathbf{p})$.

Prenant le logarithme de $L(\mathbf{p})$, la fonction du log de vraisemblance, $l(\mathbf{p})$, est alors obtenue.

$$l(\mathbf{p}) = \ln L(\mathbf{p}) = \ln K + n_{cc} \ln p_{cc} + n_{ci} \ln p_{ci} + n_{ic} \ln p_{ic} + n_{ii} \ln p_{ii} \quad (10)$$

où K est une constante réelle arbitraire.

Utilisant le fait que $p_{cc} + p_{ci} + p_{ic} + p_{ii} = 1$, l'équation 10 devient

$$l(\mathbf{p}) = \ln K + n_{cc} \ln p_{cc} + n_{ci} \ln p_{ci} + n_{ic} \ln p_{ic} + n_{ii} \ln(1 - p_{cc} - p_{ci} - p_{ic})$$

Afin de calculer les estimateurs du maximum de vraisemblance pour chaque paramètre, il suffit de résoudre la dérivée première par rapport à chaque paramètre égale à zéro ($\frac{\partial l(\mathbf{p})}{\partial p_{xy}} = 0$). Suite à cette opération nous obtenons que $\hat{p}_{cc} = n_{cc}/n$, $\hat{p}_{ci} = n_{ci}/n$

et $\hat{p}_{ic} = n_{ic}/n$.

Afin de tester l'hypothèse $H_0 : p_c = p_{.c}$ versus $H_1 : p_c < p_{.c}$, il est nécessaire de déterminer les estimateurs du maximum de vraisemblance restreints à l'hypothèse nulle. Les estimateurs (restreints et non-restreints sous H_0) serviront à bâtir le rapport de vraisemblance. Premièrement, il est important de noter que cette hypothèse nulle est vraie si et seulement si $p_{ci} = p_{ic}$. Cette conclusion est tirée directement de la définition des proportions de succès, $p_c = p_{cc} + p_{ci}$ et $p_{.c} = p_{cc} + p_{ic}$. Donc, les hypothèses nulles et alternatives deviennent $H_0 : p_{ci} = p_{ic}$ et $H_1 : p_{ci} < p_{ic}$.

Donc, selon $H_0 : p_{ci} = p_{ic}$, la fonction conditionnelle de vraisemblance devient

$$L_{H_0}(\mathbf{p}) = \frac{n!}{n_{cc}!n_{ci}!n_{ic}!n_{ii}!} p_{cc}^{n_{cc}} p_{ci}^{(n_{ci}+n_{ic})} (1 - p_{cc} - 2p_{ci})^{n_{ii}} \quad (11)$$

Utilisant les mêmes étapes que précédemment, les estimateurs sous H_0 sont

$$\begin{aligned} \hat{p}_{cc} &= n_{cc}/n \text{ et} \\ \hat{p}_{ci} &= \hat{p}_{ic} = (n_{ci} + n_{ic})/2n. \end{aligned}$$

Donc le rapport de vraisemblance est

$$T = \frac{L(\hat{\mathbf{p}})}{L_{H_0}(\hat{\mathbf{p}})} = \frac{\left(\frac{n_{cc}}{n}\right)^{n_{cc}} \left(\frac{n_{ci}}{n}\right)^{n_{ci}} \left(\frac{n_{ic}}{n}\right)^{n_{ic}}}{\left(\frac{n_{cc}}{n}\right)^{n_{cc}} \left(\frac{n_{ci}+n_{ic}}{2n}\right)^{n_{ci}} \left(\frac{n_{ci}+n_{ic}}{2n}\right)^{n_{ic}}}$$

et après quelques simplifications, T devient

$$T = 2^{n_{ci}+n_{ic}} \left(\frac{n_{ci}}{n_{ci} + n_{ic}}\right)^{n_{ci}} \left(\frac{n_{ic}}{n_{ci} + n_{ic}}\right)^{n_{ic}} \quad (12)$$

Afin de baser les tests d'hypothèses sur un point critique et étant donné la faible taille d'échantillon pour l'expérience ($n = 30$), la distribution conditionnelle exacte de T sous H_0 , étant donné $n_{c1} + n_{ic}$, doit être obtenue. Les distributions asymptotiques pour T ne seraient que des approximations peu précises [3]. De ce fait, le niveau de confiance du test serait grandement diminué. C'est pour cette raison que la première option sera envisagée.

Afin de développer la distribution conditionnelle exacte de T , il faut d'abord calculer toutes les permutations possibles de T . Pour ce faire, les valeurs n_{cc} et n_{ii} sont considérées fixes et le total des valeurs n_{c1} et n_{ic} est partagé entre les 2 positions respectives de la table de contingence (Voir l'Appendice F). Pour chacune des permutations, la statistique T est calculée. De cette façon, une distribution exacte est obtenue.

Un programme informatique a été développé afin de calculer toutes les valeurs possibles de T et de dériver la probabilité reliée à chacune. Une fois la distribution obtenue, la *p-value* est calculée directement en sommant les probabilités nécessaires reliées aux valeurs de T . A partir de cette information, une conclusion peut alors être tirée au sujet de H_0 .

Les résultats de cette analyse se trouvent à l'Appendice I.

5.2.2 Remarques

L'hypothèse $H_0 : p_c = p_{.c}$ peut être testée à l'aide de différentes statistiques. Comme précédemment décrit, un test exact peut être utilisé. Une statistique étant une approximation d'une chi-carré avec correction communément appelé le test de McNemar [2] est également convenable. Ce test a été employé par une équipe du département de périnatalogie de l'Hôpital Général d'Ottawa et sera décrit à la section I.3. Les résultats obtenues seront utilisés dans la comparaison entre les différentes techniques

statistiques choisies dans cette thèse. De plus, il est important de noter que le test de McNemar est pour tester $H_0 : p_c = p_{.c}$ versus $H_1 : p_c \neq p_{.c}$. Egalement, pour le cas où la taille de l'échantillon est considérable, le test exact et la chi-carré sont pratiquement équivalents.

5.2.3 Comparaisons multiples

La section précédente visait à comparer les techniques de contrôle 2 à la fois. L'exercice est maintenant de les comparer les 3 à la fois. L'hypothèse d'intérêt est alors de tester l'hypothèse multiple $H_0 : p_{c.} = p_{.c} = p_{..c}$ contre $H_1 : p_{c.} < p_{.c}$ ou $p_{.c} < p_{..c}$. Afin de vérifier cette hypothèse à l'aide de la méthode du maximum de vraisemblance, il faut d'abord définir la fonction de vraisemblance $L(\mathbf{p})$. Celle-ci est donnée par

$$L(\mathbf{p}) = \frac{n!}{\prod n_{xyz}!} \prod p_{xyz}^{n_{xyz}} \quad (13)$$

où $\{x, y, z \in \{c, i\}\}$.

Prenant le logarithme de $L(\mathbf{p})$, la fonction du log de vraisemblance, $l(\mathbf{p})$, est alors obtenue.

$$l(\mathbf{p}) = \ln L(\mathbf{p}) = \ln K + \sum_{x,y,z \in \{c,i\}} n_{xyz} \ln p_{xyz} \quad (14)$$

où K est une constante réelle arbitraire et $p_{iii} = 1 - p_{ccc} - p_{cci} - p_{cic} - p_{cii} - p_{icc} - p_{ici} - p_{iic}$ car $\sum_{x,y,z \in \{c,i\}} p_{xyz} = 1$.

Prenant la dérivée première de la fonction du log de vraisemblance par rapport à chaque paramètre égale à zéro ($\frac{\partial l(\mathbf{p})}{\partial p_{xyz}} = 0$). Nous obtenons alors que les estimateurs

de vraisemblance sont $\hat{p}_{ccc} = n_{ccc}/n$, $\hat{p}_{cct} = n_{cct}/n$, $\hat{p}_{ctc} = n_{ctc}/n$, $\hat{p}_{ctt} = n_{ctt}/n$,
 $\hat{p}_{icc} = n_{icc}/n$, $\hat{p}_{ict} = n_{ict}/n$ et $\hat{p}_{itc} = n_{itc}/n$.

L'hypothèse $H_0 : p_{c..} = p_{.c.} = p_{..c}$ revient à tester $H_{01} : p_{c..} = p_{.c.}$ et $H_{02} : p_{.c.} = p_{..c}$ simultanément versus $H_{11} : p_{c..} < p_{.c.}$ ou $H_{12} : p_{.c.} < p_{..c}$. De plus, ces 2 hypothèses sont également vraies si et seulement si $p_{ci.} = p_{ic.}$ et $p_{.ci} = p_{.ic}$. Encore, H_{01} et H_{02} sont vraies si et seulement si $p_{cic} = p_{icc} + p_{ict} - p_{cti}$ et $p_{cti} = p_{icc} + p_{itc} - p_{cti}$, respectivement.

Donc, la fonction de vraisemblance restreinte sous H_0 , $L_{H_0}(\mathbf{p})$, est définie en remplaçant les paramètres p_{cic} et p_{cti} par leur expression correspondante sous H_0 données précédemment. La fonction résultante est une expression complexe qui est fonction de paramètres libres du modèle. Donc, la dérivation de $L_{H_0}(\mathbf{p})$ selon chaque paramètres sera également une expression complexe. C'est pour cette raison qu'un logiciel mathématique a été utilisé afin de calculer les estimateurs de vraisemblance sous H_0 . Quelques résultats de ces calculs apparaissent à l'Appendice G.

En se basant sur la méthode suivie à la section 5.2.1, il s'avère pratiquement impossible de bâtir un test pour vérifier les hypothèses multiples à l'aide de la méthode du maximum de vraisemblance. Donc, des approches alternatives seront donc considérées subséquemment. Tout d'abord, le test de McNemar sera vu à la prochaine section.

5.3 Test de McNemar

La description sommaire de ce test est nécessaire car les résultats (obtenus des spécialistes de l'Hôpital Général d'Ottawa) sont utilisés dans la comparaison des techniques statistiques choisies. Il est important de noter la forme de la statistique afin de porter un jugement éclairé dans l'analyse.

Le test de McNemar vise à vérifier si 2 proportions de succès sont les mêmes. Spécifiquement, pour la présente expérience, les hypothèses nulle et alternative sont les suivantes.

$$H_0 : p_c = p_{.c} \text{ versus } H_0 : p_c \neq p_{.c}$$

Le but de la présente expérience est de vérifier si les 3 méthodes de surveillance foetale sont équivalentes. Elles seront comparées 2 à la fois en utilisant le test de McNemar. Le test prend la forme suivante,

$$t = \frac{(n_{.c} - n_c)^2}{n_{.c} + n_c} \quad (15)$$

où $n_{.c}$ et n_c sont le nombre de bonnes réponses de la première et de la deuxième question, respectivement.

L'hypothèse H_0 sera rejetée seulement si

$$|t| = \frac{(n_{.c} - n_c)^2}{n_{.c} + n_c} > \chi^2(1) \quad (16)$$

A noter, que le test de McNemar sert à vérifier une hypothèse bilatérale ("two-sided") en comparaison au test unilatéral basé sur le maximum de vraisemblance. De même,

ce test utilise une distribution asymptotique par opposition à une distribution conditionnelle exacte. Également, ce test est effectué à un niveau de signification α . Les résultats se trouvent à l'Appendice I.3.

La prochaine section présentera une autre alternative pour l'analyse des proportions de succès: la régression logistique. Cette technique statistique est spécialisée dans l'analyse de données discrètes.

5.4 Régression logistique

La régression est une méthode statistique utilisée pour décrire la relation entre une variable dépendante et une ou plusieurs variables indépendantes. Parfois, la variable résultante est de type discret. Depuis la dernière décennie, la régression logistique est devenue la technique de référence appliquée dans cette situation, et ce, dans plusieurs domaines [4].

Ce qui distingue la régression logistique de la régression linéaire c'est le type de la variable dépendante. Pour cette dernière, la variable est continue et pour le premier cas, la variable est discrète. Cette différence apporte un changement dans le choix et les exigences d'un modèle paramétrique à utiliser.

5.4.1 Modèle

Dans toute analyse de la régression, l'expression qui nous intéresse est la valeur moyenne de la variable dépendante (Y) étant donné la variable indépendante (x). Cette expression est dénotée par $E(Y|x)$. Avec la régression linéaire, cette expression est une équation linéaire en x , telle que $E(Y|x) = \beta_0 + \beta_1 x$, le modèle de la régression linéaire étant $Y = \beta_0 + \beta_1 x + \varepsilon$. La moyenne conditionnelle selon x de la variable dépendante, $E(Y|x)$, peut prendre n'importe quelles valeurs possibles lorsque x est entre $-\infty$ et $+\infty$. Par contre, avec une variable dépendante binaire, $E(Y|x)$ doit être comprise entre 0 et 1 (i.e. $0 \leq E(Y|x) \leq 1$). Pour la présente étude, la variable réponse ne peut prendre que 0 ou 1 comme valeurs possibles.

Si $E(Y|x)$ est non-décroissant en x , lorsque son graphique en fonction de x est tracé pour une variable discrète, la courbe qui est produite ressemble à une distribution cumulative d'une variable aléatoire et a la forme d'un S. Cette courbe montre que le changement en $E(Y|x)$ par unité de changement en x devient progressivement

plus petit lorsque $E(Y|x)$ se rapproche de 0 ou 1. De plus, la moyenne conditionnelle approche 0 et 1 “graduellement”. A partir de ces faits, il n’est pas surprenant de constater le choix de la distribution logistique pour définir la distribution de $E(Y|x)$ pour une variable Y .

Deux raisons importantes justifient le choix de la distribution logistique; d’un point de vue mathématique, elle est extrêmement flexible et facilement utilisable, et d’un point de vue biologique, son interprétation est très significative.

Afin de simplifier la notation, $\pi(x)$ sera utilisée à la place de $E(Y|x)(= \Pr(Y|x))$. La forme spécifique du modèle de la régression logistique est donc donnée par

$$Y = \pi(x) + \varepsilon \quad \text{avec } \pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \text{ et } Y \in \{0, 1\}.$$

Habituellement, $\pi(x)$ est modifiée en utilisant la transformation logit définie par

$$\begin{aligned} g(x) &= \ln \left(\frac{\Pr(Y = 1|x)}{\Pr(Y = 0|x)} \right) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) \\ &= \beta_0 + \beta_1 x \end{aligned}$$

Le groupe $Y = 1$ est comparé au groupe $Y = 0$. Donc, le groupe avec $Y = 0$ joue le rôle de groupe de référence.

La transformation logit est importante car elle permet d’obtenir les propriétés désirables du modèle de la régression linéaire. C’est-à-dire que $g(x)$ est linéaire en ses paramètres, peut être continue et peut prendre des valeurs de $-\infty$ à $+\infty$ selon les valeurs de x . De plus, $g(x)$ transforme le modèle de la régression logistique sous une forme linéaire très similaire à celle de la régression linéaire classique.

La seconde importance concerne la distribution de $Y|x$. Pour la régression linéaire, une observation de la variable dépendante est exprimée par $y = E(Y|x) + \varepsilon$. L'erreur, ε , est la différence entre une observation et la moyenne conditionnelle. L'hypothèse la plus fréquemment posée est que ε est distribuée normalement avec moyenne 0 et une variance constante. Donc, la distribution de $Y|x$ sera normale avec moyenne $E(Y|x)$ et une variance constante. Par contre, ce n'est pas le cas pour une variable dépendante binaire. Dans ce cas, une observation est donnée par $y = \pi(x) + \varepsilon$ où y a seulement 2 valeurs possibles, 0 et 1. Si $y = 1$ alors $\varepsilon = 1 - \pi(x)$ avec une probabilité de $\pi(x)$ et si $y = 0$ alors $\varepsilon = -\pi(x)$ avec une probabilité de $1 - \pi(x)$. Alors, ε suit une distribution avec moyenne 0 et une variance égale à $\pi(x)[1 - \pi(x)]$. Donc la distribution conditionnelle de la variable dépendante étant donné x est une distribution binomiale avec probabilité de succès $\pi(x)$.

Donc, après ce changement, le modèle devient

$$Y = \frac{e^{g(x)}}{1 + e^{g(x)}} + \varepsilon \text{ avec } Y \in \{0, 1\} \quad (17)$$

Ce cas avec Y binaire est toutefois le plus simple. Pour la présente expérience, nous avons des données appariées. C'est-à-dire que pour chaque sujet, une mesure a été prise pour chacun des 3 "traitements" appliqués. Le but de cette étude est de comparer ces traitements tout en tenant compte de l'appariement. Les prochaines sections présentent 2 procédés de comparaisons. Premièrement, les techniques de contrôle seront comparées 2 à la fois (comparaisons simples) et ensuite les 3 à la fois (comparaisons multiples). Ces comparaisons seront faites successivement sur les 3 variables de l'expérience, c'est-à-dire, B , P et I , et ce, avec les 6 tracés choisis pour l'expérience.

5.4.2 Comparaisons simples

Lorsque les techniques de contrôle sont comparées 2 à la fois, la variable dépendante Y prend 3 valeurs possibles; 0, 1 et 2.

Le tableau suivant montre la définition de la variable Y .

Y_1	Y_2	$Y_1 - Y_2$	Y
1	1	0	0
0	0	0	0
0	1	-1	1
1	0	1	2

Tableau 5: Valeurs prises par la variable dépendante Y lorsque 2 techniques sont comparées. Y_1 et Y_2 sont les valeurs observées pour les techniques 1 et 2, respectivement.

Donc avec un modèle contenant 3 catégories, 2 fonctions logit sont alors nécessaires. Une première pour la comparaison de $Y = 1$ versus $Y = 0$ et une deuxième pour la comparaison de $Y = 2$ versus $Y = 0$. Tout comme dans le cas binaire, le groupe avec $Y = 0$ sert de groupe de référence. Également, si l'on veut comparer $Y = 2$ à $Y = 1$, il suffit de faire la différence des logit de $Y = 2$ versus $Y = 0$ et de $Y = 1$ versus $Y = 0$. Les fonctions logit sont définies comme suit

$$\begin{aligned}g_1(x) &= \ln \left(\frac{\Pr(Y = 1|x)}{\Pr(Y = 0|x)} \right) \\ &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1p}x_p\end{aligned}$$

$$g_2(x) = \ln \left(\frac{\Pr(Y = 2|x)}{\Pr(Y = 0|x)} \right)$$

$$= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \dots + \beta_{2p}x_p$$

où les probabilités de chaque catégorie conditionnelles à la variable indépendante sont

$$\Pr(Y = 0|x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}} = \pi_0(x) \quad (18)$$

$$\Pr(Y = 1|x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} = \pi_1(x) \quad (19)$$

et finalement

$$\Pr(Y = 2|x) = \frac{e^{g_2(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} = \pi_2(x) \quad (20)$$

Par contre, pour la présente expérience, il n'y a aucune variable indépendante x . Donc, les fonctions logit sont réduites à $g_1 = \beta_{10}$ et $g_2 = \beta_{20}$.

La section suivante présente la méthode employée pour estimer les paramètres inconnus du modèle, $\beta = (\beta_{10}, \beta_{20})$. De plus, une hypothèse sera testée afin d'évaluer ces paramètres.

Hypothèse et Estimation des paramètres

Maintenant, pour développer un modèle mathématique à partir de données observées, il est nécessaire d'estimer les paramètres de ce modèle. Rappelons que le modèle de la régression logistique constitué de 3 catégories est défini par les 2 fonctions logit $g_1 = \beta_{10}$ et $g_2 = \beta_{20}$. Les paramètres inconnus à estimer sont β_{10} et β_{20} .

L'hypothèse d'intérêt avec la régression logistique lorsque la variable dépendante Y a 3 catégories et que le modèle ne contient pas de variable indépendante est donnée par

$$H_0 : \beta_{10} = \beta_{20}$$

$$H_1 : \beta_{10} \neq \beta_{20}$$

Afin de vérifier cette hypothèse, le test G est utilisé. Ce test est défini à l'aide du rapport de vraisemblance de la façon suivante,

$$G = -2 \ln \left(\frac{L(\text{modèle complet})}{L(\text{modèle réduit})} \right) \quad (21)$$

$$= -2 \ln \left(\frac{L(\hat{\beta})}{L_{H_0}(\hat{\beta})} \right) \quad (22)$$

$$= 2 \left(l_{H_0}(\hat{\beta}) - l(\hat{\beta}) \right) \quad (23)$$

Donc, les estimateurs de β pour les modèles complet et réduit doivent être calculés. Pour ce faire, supposons un échantillon aléatoire de n observations indépendantes (y_i) où y_i représente, pour le sujet i , la valeur de la variable dépendante (soit B , P ou I) prenant 0, 1 ou 2 comme valeurs possibles.

Habituellement, avec la régression linéaire, la méthode des moindres carrés est utilisée pour estimer les paramètres lorsque Y est continue. Par contre, si Y est catégorique (discrète), cette technique ne donne pas les propriétés souhaitées pour les estimateurs (non-biaisé, unique, variance minimale) [1]. Donc, la méthode d'estimation du maximum de vraisemblance (Voir la section 5.2 pour plus de détails sur l'application de la méthode) sera utilisée pour le cas où Y est de type qualitatif. Cette technique de base est d'ailleurs à l'origine de la méthode des moindres carrés. La méthode

du maximum de vraisemblance gène des valeurs pour les paramètres inconnus du modèle qui maximisent la probabilité d'obtenir les valeurs observées.

Pour construire la fonction de vraisemblance, il est nécessaire de définir 3 variables binaires Y_j ($j = 0, 1, 2$) pour indiquer l'appartenance d'une observation donnée à un groupe. Ces variables sont introduites seulement pour clarifier la fonction de vraisemblance et ne sont pas intégrées dans l'analyse de la régression logistique. Ces variables sont codifiées de la façon suivante;

Si $Y = 0$ alors $Y_0 = 1$ $Y_1 = 0$ et $Y_2 = 0$

Si $Y = 1$ alors $Y_0 = 0$ $Y_1 = 1$ et $Y_2 = 0$

Si $Y = 2$ alors $Y_0 = 0$ $Y_1 = 0$ et $Y_2 = 1$

Donc, la fonction conditionnelle de vraisemblance pour un échantillon de n observations indépendantes est

$$L(\beta) = \prod_{i=1}^n \pi_0^{y_{0i}} \pi_1^{y_{1i}} \pi_2^{y_{2i}} \quad (24)$$

Prenant le logarithme naturel et utilisant le fait que $\sum_j Y_{ji} = 1$ pour chaque i , la fonction du log de vraisemblance est

$$l(\beta) = \ln(L(\beta)) = \sum_{i=1}^n (y_{1i}\beta_{10} + y_{2i}\beta_{20}) - n \ln(1 + e^{\beta_{10}} + e^{\beta_{20}}) \quad (25)$$

La dérivation de cette fonction par rapport à β_{10} et β_{20} donne, respectivement

$$\hat{\beta}_{10} = \ln \left(\frac{\pi_1(1 + \beta_{20})}{1 - \pi_1} \right) \quad (26)$$

$$\hat{\beta}_{20} = \ln \left(\frac{\pi_2(1 + \beta_{10})}{1 - \pi_2} \right) \quad (27)$$

Maintenant, nous devons définir la fonction de vraisemblance conditionnelle à l'hypothèse nulle de l'égalité des 2 paramètres. Cette fonction est donnée par

$$l_{H_0}(\beta) = \sum_{i=1}^n (y_{1i} + y_{2i})\beta_{10} - n \ln(1 + 2e^{\beta_{10}}) \quad (28)$$

La dérivation de $l_{H_0}(\beta) = \ln(L_{H_0}(\beta))$ par rapport à β_{10} donne

$$\hat{\beta}_{10} = \ln \left(\frac{1/2(\pi_1 + \pi_2)}{1 - (\pi_1 + \pi_2)} \right) \quad (29)$$

Etant donné la non-linéarité des estimateurs (sans contrainte et avec la contrainte de H_0), des méthodes numériques itératives sont utilisées afin de calculer les estimateurs des paramètres de façon à maximiser la fonction de vraisemblance. La statistique G est par la suite comparée à une statistique χ^2_ν avec ν degrés de liberté ($\nu = \#$ paramètres du modèle complet - $\#$ paramètres du modèle réduit).

L'hypothèse $H_0 : \beta_{10} = \beta_{20}$ sera rejetée en faveur de $H_1 : \beta_{10} \neq \beta_{20}$ si G est grand. Sous H_0 , la statistique G suivra asymptotiquement une distribution χ^2 avec 1 degré de liberté ($\chi^2(1)$). Donc, la *p-value* associée à ce test est donnée par $\Pr(|G| > G_{obs} | H_0)$. Si cette probabilité est plus petite que α , alors le test est significatif.

5.4.3 Comparaisons multiples

Les comparaisons multiples visent à comparer les 3 techniques de contrôle simultanément (les 3 à la fois) (Voir le Tableau 2 au chapitre 4).

Tout comme le cas des comparaisons simples, il faut tenir compte du fait que les observations sur les 3 techniques de contrôle sont appariées. Lorsque les techniques de contrôle sont comparées les 3 à la fois, la variable dépendante Y prend 7 valeurs possibles, soit de 0 à 6. Le tableau suivant montre la définition de Y .

Y_A	1	0	1	1	1	0	0	0
Y_M	1	0	1	0	0	1	1	0
Y_V	1	0	0	1	0	1	0	1
Y	0	0	1	2	3	4	5	6

Tableau 6: Valeurs prises par la variable dépendante Y lorsque les 3 techniques sont comparées en même temps. Y_A , Y_M et Y_V sont les valeurs observées pour les techniques A, M et V, respectivement.

Donc avec un modèle contenant 7 catégories, 6 fonctions logit sont alors nécessaires. Ces fonctions logit font la comparaison de $Y = j$ versus $Y = 0$ ($j = 1, \dots, 6$). Le groupe avec $Y = 0$ sert toujours de groupe de référence. Les fonctions logit sont définies comme suit

$$\begin{aligned}
 g_j(x) &= \ln \left(\frac{\Pr(Y = j|x)}{\Pr(Y = 0|x)} \right) \\
 &= \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jp}x_p \\
 &= \beta_{j0}
 \end{aligned}$$

De façon générale, la probabilité conditionnelle d'un modèle avec 7 catégories est donnée par

$$\Pr(Y = j|x) = \frac{e^{g_j(x)}}{\sum_{k=0}^6 e^{g_k(x)}} = \pi_j(x) \quad (30)$$

où $g_0(x) = 0$ et $j = 0, \dots, 6$.

La section suivante présente l'hypothèse à tester pour le cas où le modèle contient 7 catégories et l'estimation des paramètres.

Hypothèse et Estimation des paramètres

Tout comme dans le cas des comparaisons simples, les estimateurs seront calculés selon la méthode du maximum de vraisemblance. Les paramètres inconnus à estimer sont $\beta = (\beta_{10}, \dots, \beta_{60})$.

Pour l'expérience décrite, l'hypothèse d'intérêt est donnée par

$$H_0 : \beta_{10} = \beta_{20} = 2\beta_{30} = \beta_{40} = 2\beta_{50} = 2\beta_{60} = \beta_{10}$$

opposée à

$$H_1 : \begin{array}{lll} g_1 = \beta_{10} & g_2 = \beta_{20} & g_3 = \beta_{30} \\ g_4 = \beta_{40} & g_5 = \beta_{50} & g_6 = \beta_{60} \end{array}$$

Le facteur 2 apparaissant dans l'hypothèse pour certain paramètre vient du fait que pour la valeur de Y correspondant au paramètre, une seule des variables Y_A, Y_M, Y_V égalait 1 tandis que, pour les autres paramètres sans ce facteur 2, des variables Y_A, Y_M, Y_V égalaient 1. En d'autres mots, il est 2 fois plus probable d'avoir une bonne réponse pour les 3 méthodes lorsque 2 des réponses données sont correctes que lorsqu'il n'y en a qu'une seule.

Afin de vérifier cette hypothèse, le test G défini précédemment est utilisé. Donc, les estimateurs de β pour les modèles complet et réduit doivent être calculés. Pour ce faire, supposons un échantillon aléatoire de n observations indépendantes (y_i) où y_i représente, pour le sujet i , la valeur de la variable dépendante prenant de 0 à 6

comme valeurs possibles.

Pour construire la fonction de vraisemblance, il est nécessaire de définir 7 variables binaires $Y_j = 1$ ($j = 0, \dots, 6$) si $Y = j$ pour indiquer l'appartenance à un groupe d'une observation donnée. Ces variables sont introduites seulement pour clarifier la fonction de vraisemblance et ne sont pas intégrées dans l'analyse de la régression logistique.

Donc, la fonction conditionnelle de vraisemblance pour un échantillon de n observations indépendantes est

$$L(\beta) = \prod_{i=1}^n \pi_0^{y_{0i}} \dots \pi_6^{y_{6i}} \quad (31)$$

Prenant le logarithme naturel et utilisant le fait que $\sum_j Y_{ji} = 1$ pour chaque i , la fonction du log de vraisemblance est

$$l(\beta) = \ln(L(\beta)) = \sum_{i=1}^n (y_{1i}\beta_{10} + \dots + y_{6i}\beta_{60}) - n \ln(1 + e^{\beta_{10}} + \dots + e^{\beta_{60}}) \quad (32)$$

La dérivation de cette fonction par rapport à β_{j0} , $j = 1, \dots, 6$, donne

$$\hat{\beta}_{j0} = \left(\frac{\pi_j (\sum_{k \neq j} e^{\beta_{k0}})}{1 - \pi_j} \right) \quad j = 1, \dots, 6 \quad (33)$$

Maintenant, nous devons définir la fonction de vraisemblance conditionnelle à l'hypothèse nulle. Cette fonction est donnée par

$$L_{H_0}(\beta) = \sum_{i=1}^n (y_{1i} + y_{2i} + \frac{y_{3i}}{2} + y_{4i} + \frac{y_{5i}}{2} + \frac{y_{6i}}{2}) \beta_{10} - n \ln(1 + 3e^{\beta_{10}} + 3e^{\beta_{10}/2}) \quad (34)$$

La dérivation de $l_{H_0}(\beta) = \ln(L_{H_0}(\beta))$ par rapport à β_{10} donne une expression non-linéaire. Des méthodes numériques serviront alors à évaluer cet estimateur. L'hypothèse H_0 sera rejetée en faveur de H_1 si G , la statistique développée sur le principe du maximum de vraisemblance, est grand. Sous H_0 , la statistique G suivra approximativement une distribution χ^2 avec 5 (i.e. 6 - 1) degrés de liberté ($\chi^2(5)$). Donc, la *p-value* associée à ce test est donnée par $\Pr(|G| > G_{obs}|H_0)$. Si cette probabilité est plus petite que α , alors le test est significatif.

5.4.4 Estimation

Le “rapport des chances”, $\psi_j(a, b)$, est habituellement le paramètre d'intérêt dans la régression logistique étant donné sa simplicité d'interprétation. Ce facteur donne une mesure d'association dans le sens qu'il donne une approximation de la probabilité du résultat d'être présent parmi les individus avec $x = a$ par rapport à ceux avec $x = b$. Il est défini comme suit

$$\psi_j(a, b) = \frac{\Pr(Y = j|x = a)/\Pr(Y = 0|x = a)}{\Pr(Y = j|x = b)/\Pr(Y = 0|x = b)} \quad (35)$$

Toutefois, son estimateur, $\hat{\psi}_j(a, b)$, tend à avoir une distribution biaisée. En théorie, $\hat{\psi}_j$ a une distribution normale pour de grands échantillons. Mais cette exigence sur la taille de l'échantillon est rarement satisfaite pour la plupart des études. Donc, les inférences sont généralement faites sur $\ln(\hat{\psi}_j(a, b))$ qui tend à suivre une distribution normale, et ce, avec une exigence réduite sur la taille de l'échantillon.

Un intervalle de confiance à $100(1-\alpha)\%$ de confiance pour le “rapport des chances” est obtenu en calculant premièrement les valeurs d'un intervalle de confiance pour le paramètre β_{ij} . Par la suite, ces valeurs sont passées en arguments à la fonction exponentielle pour finalement obtenir

$$\psi_j(a, b) \in \left[\exp \left\{ \hat{\beta}_{1,j} \pm Z_{(1-\alpha/2)} \sqrt{\frac{\widehat{\text{Var}}(d_{1,j})}{n}} \right\} \right] \quad (36)$$

Pour la présente étude, le “rapport de chances” ne peut être calculé car l’analyse est faite sans variable indépendante.

La section suivante apporte quelques remarques au sujet de la régression logistique appliquée à la présente expérience.

5.4.5 Remarques

La technique de la régression logistique définie précédemment utilise une approche semblable à celle utilisée à la section 5.2, la méthode du maximum de vraisemblance. Par contre, avec la présente technique, les points critiques de la distribution χ^2 sont employés.

Le logiciel SAS possède une procédure spécifique pour la régression logistique appelée PROC LOGISTIC. Après exécution de cette procédure, il semble que cette approche ne soit pas adéquate pour tester l’ensemble de données. Il existait des cas presque évident de données où l’hypothèse nulle n’aurait pas dû être rejetée et le test basé sur un point critique χ^2 ne l’acceptait pas. Comme par exemple, considérons le cas de 3 catégories. La catégorie de référence comptait un effectif de 15 unités et les 2 autres avaient respectivement 7 et 8 unités. Il semble clair que l’hypothèse de l’égalité des 2 proportions de ces catégories ne devrait pas être rejetée car elles ont presque le même effectif.

De plus, SAS semble avoir quelques défaillances au niveau des options passées à la procédure LOGISTIC. Premièrement, il était impossible de spécifier au programme

le nombre total de catégories qu'un ensemble de données devait avoir. Cela aurait été utile lorsque les données étudiées ne satisfaisaient pas 1 ou plusieurs catégories. Il a fallu ajuster les bases de données pour répondre à ce besoin afin que le test soit conforme à ce qui a été défini.

Analyse Multivariée

Une analyse multivariée dans ce cas-ci correspond à inclure les 3 variables de l'étude dans l'analyse. Ces variables sont B , P et I . Mais en ce qui concerne la régression logistique dans une situation où les observations sont appariées et que la taille de l'échantillon est limitée à 30, une analyse multivariée semble non réalisable.

Supposons qu'une analyse multivariée était considérée, les vecteurs des variables auraient une dimension de 3×1 (Voir la section 4 pour la définition des vecteurs.). Dans le cas d'une analyse univariée, un test a été développé nécessitant la création d'une variable Y de 3 catégories lorsqu'une hypothèse simple était vérifiée et de 7 catégories pour le cas d'une hypothèse multiple. Si un test était développé pour analyse multivariée, il serait nécessaire de définir une variable Y de $3 \times 3 \times 3 (=27)$ catégories pour tester une hypothèse simple et de $7 \times 7 \times 7 (=343)$ catégories pour tester une hypothèse multiple.

A ce point, il est important de rappeler que l'échantillon de la présente expérience n'est que de 30 sujets. Donc, si le nombre de catégories de la variable étudiée est plus élevée que le nombre de sujets, alors il ne semble pas logique de développer un tel test lorsque n est petit.

Etant donné cette contrainte, ceci nous amène à envisager l'utilisation d'une autre méthode statistique disponible; le ré-échantillonnage. Cette technique semble la plus adaptée et la moins contraignante pour l'expérience donnée. La prochaine partie expliquera en quoi consiste cette technique.

5.5 Ré-échantillonnage

Le terme "ré-échantillonnage" fait référence à des méthodes avec lesquelles les valeurs observées sont utilisées intensivement. Habituellement, ces manipulations sont faites à l'aide d'ordinateurs. Le ré-échantillonnage fait ce qu'un chercheur ferait en pratique, c'est-à-dire, il répèterait l'expérience plus d'une fois, et ce, des milliers de fois [7].

Le ré-échantillonnage vise ré-assigner aléatoirement les données expérimentales parmi les unités de l'échantillon. A partir de comparaisons faites entre les valeurs observées et les valeurs ré-échantillonnées, nous pouvons conclure si l'échantillon est *INHABITUEL* (vraiment significatif) ou *TYPIQUE* (dû à la chance).

Pour l'expérience décrite au chapitre 2, considérons les vecteurs Y_1, \dots, Y_n dénotant les observations indépendantes ayant une distribution multivariée binomiale $(\pi, 1)$. Chacune de ces observations a une dimension 3. C'est-à-dire qu'une mesure a été enregistrée pour chacune des méthodes de contrôle de la condition foetale. Donc, les observations sont $Y_i = [Y_{i1}, Y_{i2}, Y_{i3}]$ ($i = 1, \dots, 30$). Les Y_i sont des variables indépendantes. Pour chaque i , les Y_{ij} ($j = 1, 2, 3$) sont binomiales $(\pi_i, 1)$ mais elles peuvent dépendre l'une de l'autre. Le type de cette dépendance est cependant inconnue.

Le vecteur des fréquences binomiales observées est dénoté par $B = [B_j] = \sum_{i=1}^{n=30} Y_{ij}$ ($j = 1, 2, 3$) et a une distribution multivariée binomiale avec probabilités π . De plus, les éléments individuels B_j sont distribués binomialement avec paramètres $(\pi_j, n = 30)$.

Les hypothèses d'intérêt présentées au chapitre 4 seront vérifiées à l'aide des intervalles de confiance simultanés de Clopper-Pearson. Ces hypothèses ont la forme $H_{0j} : \pi_j = \pi_{0j}$, pour une constante fixée π_{0j} . Les intervalles de Clopper-Pearson pour π_j ont été choisis plutôt que les intervalles basés sur la distribution normale car des intervalles symétriques venant de la théorie normale sont moins désirables pour des proportions près de 0. Ces intervalles sont parfois caractérisés comme étant "exacts" mais ils ne le sont pas vraiment étant donné la nature discrète de la distribution binomiale ([7], page 178). Le fait que les intervalles de Clopper-Pearson soient basés sur une densité exacte est une autre raison pourquoi on les préfère aux intervalles basés sur une distribution normale.

Les intervalles de confiance de Clopper-Pearson sont maintenant définis. Notons que chaque B_j ($j = 1, 2, 3$) est distribué comme $BIN(\pi_j, n)$. Si la valeur observée est b_j , alors la borne supérieure de Clopper-Pearson est la valeur $\pi_S(b_j)$ telle que

$$\Pr(B_j \leq b_j | \pi_S(b_j)) = \sum_{i=0}^{b_j} \binom{b_j}{i} \pi_S^i (1 - \pi_S)^{n-i} = \alpha/2 \quad (37)$$

Similairement, la borne inférieure $\pi_I(b_j)$ est définie telle que

$$\Pr(B_j \geq b_j | \pi_I(b_j)) = \sum_{i=b_j}^n \binom{n}{i} \pi_I^i (1 - \pi_I)^{n-i} = \alpha/2 \quad (38)$$

Alors l'intervalle de confiance marginal (non simultané) de Clopper-Pearson à un niveau de $100(1-\alpha)\%$ est l'ensemble de valeurs $I_\alpha(b_j)$ tel que

$$I_\alpha(b_j) = \begin{cases} [0, \pi_S(b_j)) & \text{pour } b_j = 0 \\ (\pi_I(b_j), \pi_S(b_j)) & \text{pour } 0 < b_j < n \\ (\pi_I(b_j), 1] & \text{pour } b_j = n. \end{cases}$$

Les équations 37 et 38 sont résolues numériquement. Etant donné que \mathbf{B} est en fait un vecteur de v observations ($v = 3$ mesures répétées), l'intervalle $I_{j\alpha} = I_\alpha(b_j)$ doit être calculé pour chaque $j = 1, \dots, v$. Le niveau de signification marginal de chaque intervalle est approximativement $100(1 - \alpha)\%$. Par contre, pour une inférence simultanée, le paramètre d'intérêt est

$$C_\alpha = \Pr(\pi_j \in I_\alpha(B_j), \text{ pour } j = 1, 2, 3) \quad (39)$$

Cette donnée ne peut être obtenue facilement de façon analytique mais peut être estimée en utilisant le ré-échantillonnage *bootstrap*. Définissons Y_1^*, \dots, Y_n^* comme étant un échantillon *bootstrap* obtenu à partir des vecteurs originaux Y_1, \dots, Y_n sous l'hypothèse nulle. Ceci sera décrit plus tard. Egalement, dénotons les fréquences *bootstrap* calculées à partir des nouvelles données par B_j^* . Ainsi, la probabilité de couverture simultanée C_α peut être estimée par

$$\hat{C}_\alpha = \Pr(B_j^*/n \in I_\alpha(B_j^*), \text{ pour } j = 1, 2, 3) \quad (40)$$

utilisant le ré-échantillonnage *bootstrap*.

La valeur \hat{C}_α est relativement simple à obtenir pour un α fixe. Premièrement, il faut créer les intervalles de Clopper-Pearson pour chaque $j = 0, \dots, n$ avant le ré-échantillonnage. Les intervalles de confiance pour $n = 30$ avec $\alpha = 0.05$ se trouvent à l'Appendice H. Par la suite, il suffit d'identifier les intervalles $I_\alpha(B_j^*)$ à partir de la liste d'intervalles pré-établie pour l'échantillon *bootstrap*. Si les proportions B_j^*/n

sont contenues dans l'intervalle $I_\alpha(Y_j^*)$, alors un compteur est incrémenté. La valeur estimée \hat{C}_α est alors la proportion de simulations pour lesquelles toutes les B_j^*/n sont contenues dans l'intervalle $I_\alpha(B_j^*)$.

Il est important à ce point d'expliquer brièvement en quoi consiste le ré-échantillonnage *bootstrap* en général et comment il est appliqué dans le cas présent. La méthode du *bootstrap* est une technique de ré-échantillonnage dite avec remplacement. Cette technique consiste à considérer les observations initiales Y_1, \dots, Y_n comme une population fixe de nombres. Les observations ré-assignées Y_i^* sont générées aléatoirement et indépendamment comme suit:

$$Y_i^* = \begin{cases} Y_1, \text{ avec probabilité } 1/n \\ Y_2, \text{ avec probabilité } 1/n \\ \vdots \\ Y_n, \text{ avec probabilité } 1/n \end{cases}$$

La valeur Y_i^* ($i = 1, \dots, n$) sélectionnée est ensuite retournée parmi les Y_1, \dots, Y_n . Il est ainsi possible qu'une observation soit présente plus d'une fois car le ré-échantillonnage *bootstrap* est fait avec remise. Avec le ré-échantillonnage sans remplacement, les observations reviennent qu'une seule fois. Elles ne sont que permutées entre elles.

De cette façon, il est possible d'estimer la variance des (B_j/n) . Par contre, ceci n'est pas suffisant si l'on veut estimer C_α . La valeur C_α désigne la *p-value* associée à B . C'est-à-dire, la probabilité que B est *typique* quand H_0 est vraie. Donc, il faut ré-échantillonner sous la contrainte de H_0 . Choisir Y_1^*, \dots, Y_n^* aléatoirement de $\{Y_1, \dots, Y_n\}$ ne donnera pas un échantillon *bootstrap* où l'effet de H_0 est incorporé. Donc, la valeur C_α estimée ne reflètera pas si la variable B est *typique* ou *inhabituelle* sous H_0 .

Si H_0 est vraie, les 3 façons d'observer les tracés sont donc équivalentes. Donc, si

par exemple, $Y_1 = (a, b, c)$ avec $(a, b, c) \in \{0, 1\}$, H_0 veut dire que les permutations de Y_1 ; $Y_1^{(1)} = (a, b, c)$, $Y_1^{(2)} = (a, c, b)$, $Y_1^{(3)} = (b, a, c)$, $Y_1^{(4)} = (b, c, a)$, $Y_1^{(5)} = (c, a, b)$ et $Y_1^{(6)} = (c, b, a)$ sont aussi probables. Chaque permutation des résultats à l'intérieur d'un vecteur Y_i est raisonnable sous H_0 . En effet, chaque sujet est traité comme un "bloc" et les observations sont perméutées dans ce bloc. Donc, le ré-échantillonnage *bootstrap* sous H_0 est de choisir $\{Y_1^* \dots Y_n^*\}$ (avec remise) de

$$\cup_{k=1}^6 \{Y_1^{(k)}, Y_2^{(k)}, \dots, Y_n^{(k)}\}$$

où $Y_i^{(k)}$ représente la $k^{\text{ième}}$ permutation des 3 observations d'un bloc. Cet échantillonnage est répété des milliers de fois et C_α est calculée. Pour le cas présent, 5000 simulations seront effectuées.

Le chapitre suivant présente l'analyse des différentes méthodes statistiques précédemment décrites. Ces techniques étaient celles du maximum de vraisemblance, de la régression logistique et du ré-échantillonnage. De plus, les résultats obtenus avec ces techniques seront comparés à ceux obtenus avec le test de McNemar. Ce test a été employé par une équipe de l'Hôpital Général d'Ottawa pour spécifiquement vérifier l'équivalence des 3 différentes techniques de contrôle de la surveillance foetale.

Chapitre 6

Analyse des résultats et comparaisons des méthodes statistiques

Etant donné que le test de Student n'était pas valable pour l'expérience décrite, la comparaison des techniques statistiques ne comprendra que la méthode du maximum de vraisemblance, la régression logistique et le ré-échantillonnage. La section suivante comparera ces techniques lorsque des comparaisons simples sont vérifiées.

6.1 Comparaisons simples

En ce qui concerne la régression logistique, il semble qu'une valeur critique basée sur la distribution χ^2 ne soit pas efficace à détecter les différences entre les proportions de succès des méthodes de contrôle. Sur 36 tests effectués, 11 résultent à une *p-value* de 1.0000. Cette valeur est obtenue quand le log de vraisemblance est le même pour les modèles réduit et complet. Ce fait survient lorsqu'une des catégories de la variable aléatoire est manquante. Alors, lorsque c'est le cas, cela veut dire que l'effectif de cette catégorie est 0. Donc, la probabilité de succès est nulle et l'on peut facilement

conclure que les probabilités de succès sont différentes car une est 0 et l'autre diffère de 0. Le test de la régression logistique dans ces cas donne des probabilités égales. Ceci doit être considéré faux.

Si l'on compare les cas où les résultats de la régression logistique donnant une *p-value* de 1.000 avec les *p-values* obtenues avec le test de McNemar, ceux-ci conclut effectivement que les probabilités de succès sont différentes. Il semble que lorsque la différence entre les effectifs est plus grande que 5, le test de McNemar conclut qu'il y a une différence entre les techniques. Seulement un cas sur 11 conclut à une égalité (il y a seulement une différence de 1 entre les effectifs pour les 2 techniques étudiées).

Pour ces mêmes cas, avec le test basé sur le maximum de vraisemblance, des conclusions semblables à celles basées sur le test de McNemar sont obtenues. Avec la méthode du maximum de vraisemblance, le test avait été bâti de façon que l'on pouvait conclure si $p(A)$ était plus petit que $p(M)$ lorsque les méthodes A et M étaient comparées ou bien conclure si $p(M)$ était plus petit que $p(V)$ lorsque les méthodes M et V étaient comparées. Donc, ici les conclusions ont cette forme plutôt que de conclure que les techniques sont différentes (\neq).

Pour les autres cas (*p-values* \neq 1.0000), la régression logistique semble donner des conclusions exactes que dans 39% des cas (14 cas sur 36). Avec ce pourcentage de succès, il ne paraît pas possible d'affirmer que le test basé sur la régression logistique soit efficace lorsqu'il est appliquée à une telle expérience.

Il est également important de noter que le test basé sur la régression logistique est obtenu à l'aide de la méthode du maximum de vraisemblance mais la décision sur H_0 ou H_1 est basée sur un point critique de la densité χ^2 plutôt que sur une distribution exacte. Si la taille de l'échantillon avait été plus grande, il est plus probable que le test ait été plus sensible aux variations de l'échantillon.

Les tests de McNemar et du maximum de vraisemblance aboutissent aux mêmes

conclusions sur l'égalité des probabilités de succès. De plus, lorsque le test de McNemar conclut que $p(A) \neq p(M)$ ou que $p(M) \neq p(V)$ alors le test basé sur le maximum de vraisemblance concluait que $p(A) < p(M)$ ou que $p(M) < p(V)$ car le test permettait d'arriver à ce type de conclusion. Egalement, pour les 2 tests, les *p-values* obtenues étaient très semblables.

Par contre, dans un des cas, le test de McNemar accepte l'hypothèse nulle tandis que le test du maximum de vraisemblance la rejette. Mais si les *p-values* étaient précises qu'à deux décimales près, les 2 tests accepteraient H_0 . Egalement, dans 2 autres cas, les conclusions sont différentes. Le test de McNemar conclut que les probabilités sont différentes tandis que le test du maximum de vraisemblance conclut à leur égalité. Ces 2 conclusions sont correctes car pour ces 2 cas, il est possible de voir qu'en fait $p(M)$ n'est pas plus petit que $p(V)$ mais bien que $p(M) > p(V)$. Donc, il est correct d'accepter l'hypothèse nulle dans le cas du maximum de vraisemblance.

Les résultats tirés avec la méthode du maximum de vraisemblance seront utilisés comme référence pour comparer les résultats obtenus du ré-échantillonnage. Premièrement, après un examen détaillé des données brutes, il semble que les résultats obtenus de ce test décrivent bien l'échantillon tiré. En fait, H_0 était rejetée lorsqu'il y avait une différence significative entre les probabilités de succès de 2 méthodes de contrôle. Egalement, étant donné qu'un deuxième test, le test de McNemar, rapporte les mêmes conclusions, ceci renforce davantage la confiance sur le test du maximum de vraisemblance.

Pour ce qui est du ré-échantillonnage, 5000 nouveaux échantillons *bootstrap* ont été tirés à l'aide de cette technique. Lorsque les résultats sont comparés à ceux du maximum de vraisemblance, il s'avère que les conclusions sont relativement les mêmes à l'exception de quelques cas. Parmi ceux-ci, il semble que le facteur commun est le faible nombre de personnes ayant donné des réponses correctes à précisément une des 2 méthodes étudiées. Cette valeur est la quantité $n_{ca} + n_{ic}$ considérée avec la méthode du maximum de vraisemblance. Si une unité était soit soustraite ou additionnée à

ce nombre, les conclusions obtenues des 2 méthodes coïncideraient. De fait, lorsque $n_{ci} + n_{ic}$ est petit, la *p-value* change très vite avec n_{ci} (ou n_{ic}). Souvent, il ne suffit que d'une seule unité pour que la conclusion soit complètement différente. Parmi les 6 tracés, 8 de ces cas existent sur 36 (6 tests pour chacun des 6 tracés).

Il semble que pour tester des comparaisons simples, la méthode du maximum de vraisemblance reflète bien ce qui existe dans l'échantillon. C'est-à-dire qu'il nous paraît efficace à détecter les différences entre les paramètres. Également, le test de McNemar semble au même niveau par rapport aux performances du test. Par contre, la statistique basée sur la régression logistique n'est pas conseillée car elle ne semble pas sensible aux variations des paramètres. Finalement, le ré-échantillonnage semble également être efficace mais il faut vérifier la cohérence des résultats lorsque $n_{ci} + n_{ic} < 10$.

La section suivante discutera maintenant de comparaisons multiples testées à l'aide de 2 des 3 techniques statistiques choisies.

6.2 Comparaisons multiples

L'étude des techniques statistiques dans le cadre des comparaisons multiples sera faite seulement avec les techniques de la régression logistique et du ré-échantillonnage. Il a été possible de créer un test pour les comparaisons multiples qu'avec ces méthodes.

Pour 8 cas sur 18 (3 tests pour chacun des 6 tracés), les conclusions tirées à la suite de l'application de la régression logistique et du ré-échantillonnage sont complètement opposées. En fait, les incohérences obtenues avec les comparaisons simples au sujet de la régression logistique ne sont que répercutées ici.

De plus, dans 2 cas, l'hypothèse nulle a été acceptée avec le ré-échantillonnage

et lorsque les résultats des 2 comparaisons simples constituant l'hypothèse multiple sont vérifiées, H_0 aurait dû être rejetée. C'est plus difficile de rejeter H_0 dans un contexte multiple.

6.3 Remarques

Pour l'expérience étudiée, il semble que des tests pour une analyse multivariée soit relativement difficile à développer. Des tests basés sur le ré-échantillonnage ont les meilleures chances de réussir, mais la petite taille de notre échantillon a limité son efficacité.

Finalement, il est important de noter que ce document n'avait que des considérations statistiques. Malgré ce fait, il faut toujours tenir compte du contexte d'une expérience lors de l'analyse des données. Ici, aucune considération clinique n'a été faite au sujet des techniques de contrôle de la surveillance foetale. De même, cette thèse ne dit rien sur l'efficacité des 3 techniques de surveillance.

6.4 Conclusions

Lorsqu'une analyse est considérée, il faut toujours apporter une attention particulière au choix d'une statistique pour tester une hypothèse donnée. Il faut tenir compte du type de données étudiées de même que des conditions requises pour l'application d'un test. Également, comme il a été démontré dans le présent document, il existe parfois une seule technique qui est vraiment efficace pour l'étude d'un sujet.

De plus, cette thèse a permis de faire en quelque sorte un bref historique des techniques statistiques utilisées depuis une cinquantaine d'années. Le but premier était

purement statistique; c'est-à-dire qu'il visait l'évaluation des techniques d'analyse choisies dans le cadre d'une expérience médicale. Ce but a été atteint pour le cas d'une analyse univariée mais n'a pu être possible pour le cas multivarié.

Les pages qui suivent présentent de l'information complémentaire pour les lecteurs intéressés à en savoir davantage au sujet de la logistique entourant cette thèse.

Annexe A

Description technique du matériel employé

Cette section présente les différents logiciels et ordinateurs utilisés pour effectuer certaines tâches nécessaires à l'analyse des données.

Avant tout, il est important de décrire brièvement le logiciel utilisé lors de l'expérience servant à présenter les tracés. Ce logiciel a été développé spécialement pour cette expérience par Virtual Medicine Corporation d'Ottawa et porte le nom de TRACER. L'ordinateur sur lequel le logiciel était installé était équipé d'une carte graphique et de haut-parleurs afin de reproduire le son enregistré avec un stéthoscope et afin de présenter graphiquement les tracés.

La première méthode consiste à compter le nombre de battements cardiaques foetaux par minute à l'aide d'un stéthoscope. Donc, le programme fait entendre les battements de coeur pendant 3 minutes. Avec cette information auditive, le sujet doit calculer les battements du coeur foetal pour des périodes de 15 secondes; soit avant, pendant et après la contraction. Pendant la période de 3 minutes, une barre verticale montre l'intensité de la contraction. A la fin de chaque représentation de ce type, le candidat doit convertir ses calculs en battements par minute (bpm) un utilisant une table de multiplication afin d'éviter les erreurs de calcul.

La deuxième technique est présentée de façon similaire à la première à l'exception qu'un compteur digital affiche à l'écran le calcul des battements cardiaques. Cette information est mise à jour toutes les 2 secondes. Le candidat n'a alors pas à compter les battements mais il peut toutefois les entendre. Egalement, tout comme la première méthode, l'intensité des contractions est présentée à l'écran.

Finalement, en ce qui trait à la technique visuelle, les tracés montrant les battements cardiaques et la contraction sont présentés à l'écran. Pendant la période de 3 minutes, une barre verticale se déplace sur les tracés pour montrer où s'appliquent les battements entendus. Egalement, le compteur digital affiche le nombre des battements cardiaques et l'intensité de la contraction est aussi présentée.

Le logiciel SAS (SAS Institute Inc.) (version 6) a servi à accomplir la majeure partie de l'analyse statistique. Ce logiciel était installé sur un système RS/6000 de IBM régi par le système d'exploitation AIX (UNIX). Egalement installé sur ce système, le logiciel MATHEMATICA a permis de faire la résolution de problèmes numériques.

De plus, un programme informatique a été développé en Borland PASCAL (version 7) pour faciliter les calculs nécessaires reliés à la méthode du maximum de vraisemblance. Egalement, le logiciel RESAMPLING STATS (version 3.0) a servi à effectuer certaines tâches relatives au ré-échantillonnage. Ces logiciels étaient installés sur un micro-ordinateur Pentium de marque Dell exploité sous Windows 95.

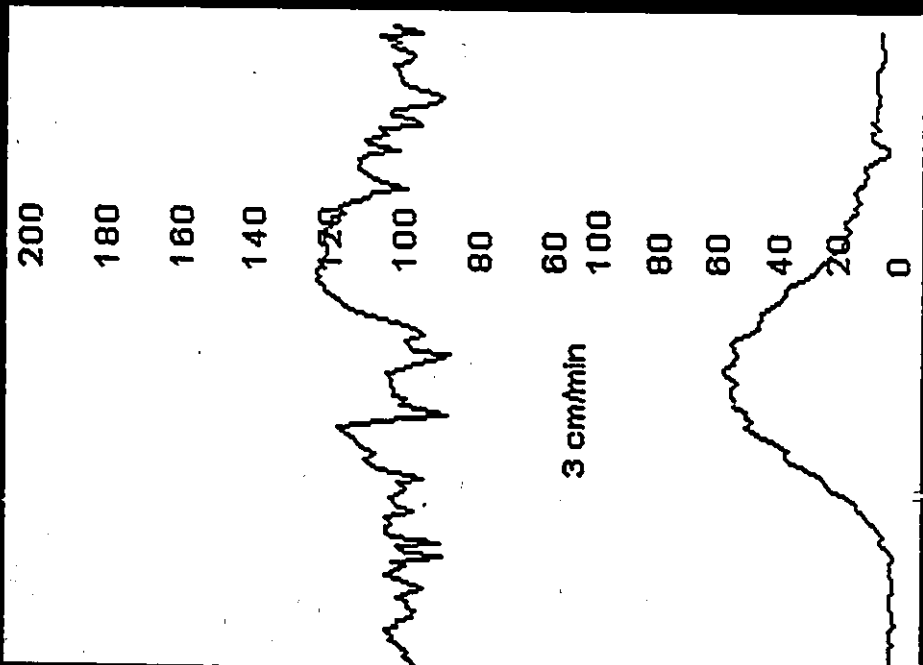
Annexe B

Présentation des tracés

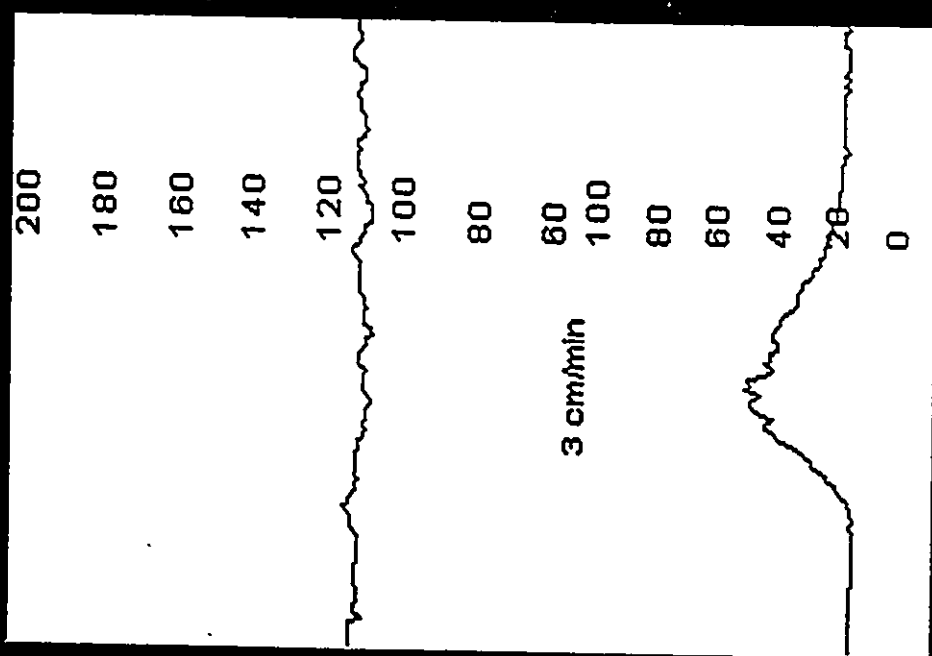
Les 3 prochaines pages présentent les 6 tracés ayant servi à tester les méthodes d'évaluation de la condition foetale qui étaient

1. le diagnostic avec un stéthoscope appuyé sur le ventre de la mère (compter les battements du coeur foetal)
2. le diagnostic avec un stéthoscope aidé d'un compteur électronique des battements du coeur foetal
3. le diagnostic avec une représentation graphique des courbes des battements du coeur foetal et d'une contraction

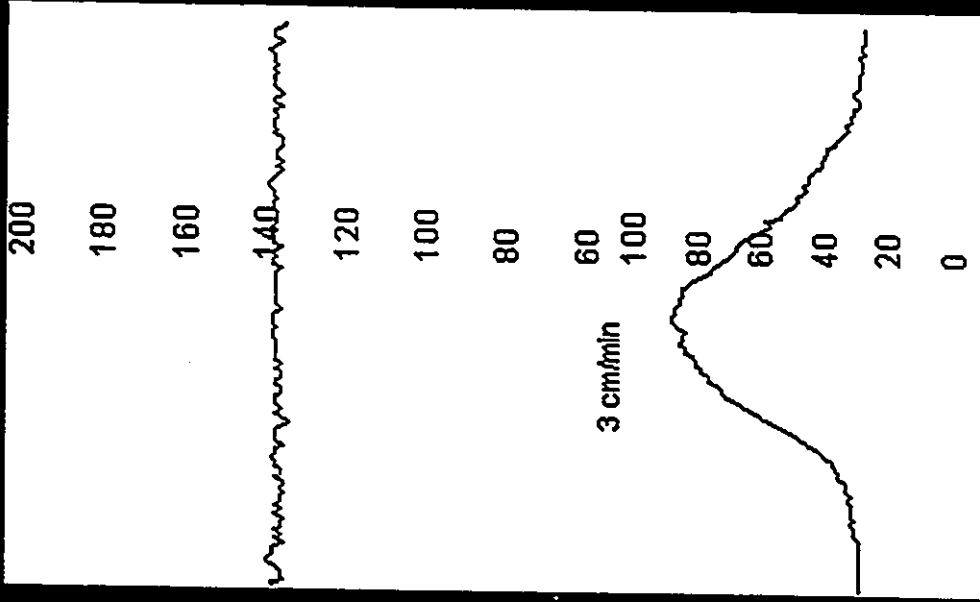
La courbe du haut de chaque graphique représente le taux des battements cardiaques foetaux mesurés en battements par minute (bpm). Ce tracé sert à répondre aux questions 1 et 2 du questionnaire (le taux et la variabilité de base des battements cardiaques). La courbe du bas est appelée le TOCO et est une mesure de la pression exercée par l'utérus et le pic de cette courbe dénote une contraction. L'analyse de cette courbe, combinée à la première, permet de répondre aux questions 3 et 4 (changements périodiques des battements de coeur et interprétation). Ces informations sont enregistrées à l'aide d'un moniteur attaché à l'aide d'une bande élastique sur le ventre de la mère.



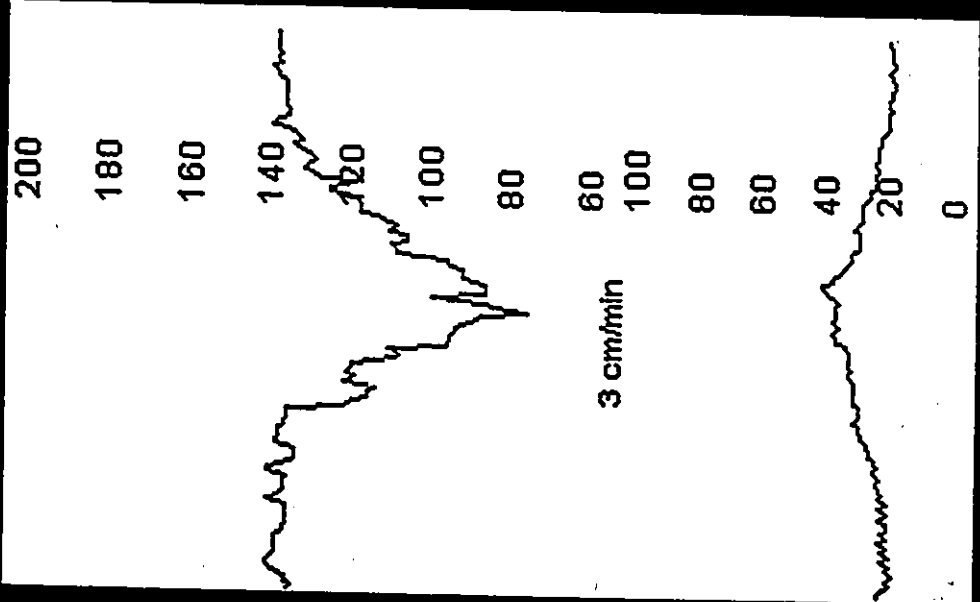
1



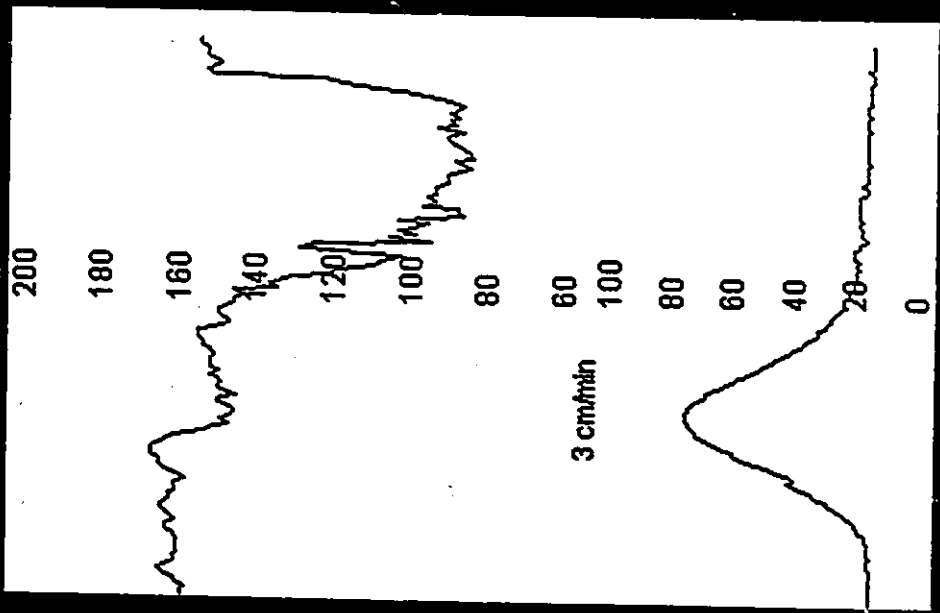
2



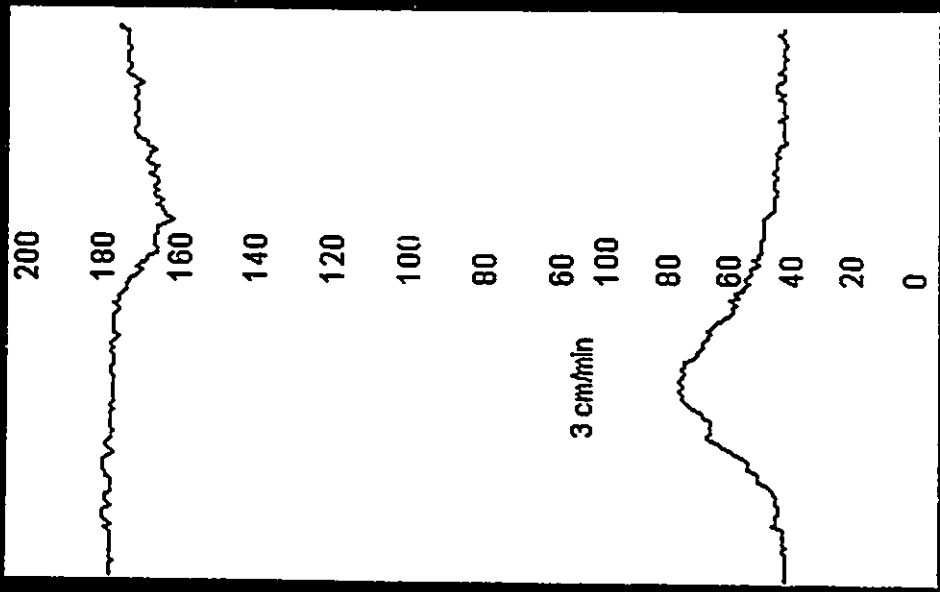
4



3



5



6

Annexe C

Exemplaire du questionnaire utilisé pour l'expérience

La prochaine page présente le questionnaire utilisé lors de l'expérience servant à évaluer les 3 techniques de contrôle de la surveillance foetale.

TRIALS IN FETAL SURVEILLANCE								
		1	2	3	...	16	17	18
BASELINE RATE (bpm) - please record ...								
Question 1								
BASELINE VARIABILITY	cannot assess? (*)							
	reduced? (< 6 bpm) (-1)							
	normal? (6-25 bpm) (0)							
	excessive? (> 25 bpm) (1)							
Question 2								
PERIODIC CHANGES	cannot assess? (*)							
	not present? (-2)							
	acceleration? (-1)							
	"early" deceleration? (1) (i.e. recovery by end of contraction)							
	"late" deceleration? (2) (i.e. recovery after end of contraction)							
	please record lowest heartrate in bpm...							
Question 3								
INTERPRETATION	reassuring? (0)							
	non-reassuring? (1)							
Question 4								

Tableau 1: Grille utilisée pour enregistrer les données lors de l'expérience sur la surveillance foetale. Les résultats ont été compilés pour les 18 versions des tracés produits à partir de 6 tracés présentés selon les 3 différentes modalités. Seulement les items codifiés (ceux avec valeurs entre parenthèses) ont été utilisés pour l'étude.

Annexe D

Liste des données originales

Les pages qui suivent donnent les résultats du questionnaire présenté à 30 résidents et infirmières du département d'obstétrique et de gynécologie de l'Hôpital Général d'Ottawa concernant les 6 tracés choisis pour l'expérience.

DONNEES POUR LA TRACE 1									
Sujet #	A_B	A_P	A_I	M_B	M_P	M_I	V_B	V_P	V_I
1	0	0	0	1	0	1	1	1	1
2	0	0	0	1	1	1	1	1	1
3	0	1	1	1	0	1	0	1	1
4	0	0	0	1	0	0	1	1	1
5	0	0	1	1	1	0	1	1	1
6	0	0	0	0	1	0	1	1	1
7	0	0	0	1	1	1	1	1	1
8	0	0	0	1	1	1	1	1	1
9	1	1	1	1	1	0	1	1	1
10	1	1	1	1	0	0	0	1	0
11	0	0	0	1	1	0	1	1	1
12	0	1	0	1	1	1	1	1	1
13	1	0	0	1	1	0	1	1	1
14	0	1	0	1	1	1	1	1	1
15	0	0	0	0	1	1	1	1	1
16	0	1	1	0	1	0	1	1	1
17	0	0	0	1	1	0	1	1	0
18	0	0	0	1	0	0	1	1	0
19	0	0	0	1	0	0	1	0	1
20	0	1	0	1	0	1	1	1	1
21	0	1	0	1	1	0	1	1	0
22	0	1	1	0	1	0	1	1	1
23	1	0	1	1	0	0	1	1	1
24	0	0	0	0	1	0	1	1	1
25	1	0	0	1	1	0	1	1	0
26	1	1	1	1	1	1	1	1	1
27	0	0	0	0	1	0	1	1	1
28	1	0	1	1	1	0	1	1	1
29	0	0	0	1	1	0	1	1	1
30	0	1	0	1	1	0	1	1	1

Tableau 2: Résultats du questionnaire concernant la Trace 1. Les techniques sont dénotées par A, M et V et les variables mesurées le sont par B, P et I.

DONNEES POUR LA TRACE 2									
Sujet #	A_B	A_P	A_I	M_B	M_P	M_I	V_B	V_P	V_I
1	1	0	1	1	1	1	1	1	1
2	0	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	0	0	0
4	0	0	0	1	0	0	1	0	1
5	0	1	1	0	0	0	1	1	1
6	1	0	0	1	1	0	1	1	1
7	0	0	0	1	0	1	0	1	0
8	1	0	1	1	0	0	1	0	0
9	0	1	0	1	1	1	1	1	1
10	0	1	0	1	1	1	0	1	0
11	0	0	1	1	0	1	1	0	1
12	1	0	1	1	1	1	0	1	1
13	1	0	1	1	0	1	1	1	1
14	0	0	0	1	0	1	1	1	1
15	0	1	0	0	1	0	1	1	0
16	0	0	1	0	1	0	1	1	0
17	1	0	1	1	1	1	1	1	1
18	1	0	1	1	0	1	0	1	1
19	0	1	1	1	0	0	1	1	0
20	0	0	0	1	0	1	0	0	1
21	0	1	1	1	1	1	1	1	1
22	0	0	1	1	1	1	1	1	1
23	1	0	1	1	0	1	1	1	1
24	0	0	1	1	1	1	1	1	1
25	1	0	1	1	1	1	1	1	1
26	0	0	1	1	1	1	1	1	1
27	0	1	0	0	1	0	0	1	0
28	0	0	1	1	1	1	1	1	1
29	0	0	0	1	1	1	1	1	1
30	0	1	0	1	0	1	1	1	1

Tableau 3: Résultats du questionnaire concernant la Trace 2. Les techniques sont dénotées par A, M et V et les variables mesurées le sont par B, P et I.

DONNEES POUR LA TRACE 3									
Sujet #	A _B	A _P	A _I	M _B	M _P	M _I	V _B	V _P	V _I
1	1	1	1	1	1	1	1	1	1
2	0	1	0	1	1	1	0	1	0
3	1	1	1	1	1	1	0	1	0
4	0	1	1	1	1	0	0	1	0
5	0	1	0	0	1	0	0	1	0
6	0	1	0	0	1	1	0	1	0
7	1	1	1	1	1	1	0	1	0
8	1	1	1	1	1	0	1	1	1
9	1	1	0	1	1	1	1	1	1
10	0	1	0	0	1	0	0	1	0
11	1	1	1	1	1	1	0	1	0
12	0	1	1	0	0	0	1	1	0
13	1	1	0	0	1	0	0	1	0
14	1	1	1	0	1	0	0	1	0
15	0	1	1	1	1	1	0	1	1
16	0	1	0	0	1	1	0	1	0
17	1	0	0	1	0	0	1	1	0
18	0	0	0	1	1	1	0	1	0
19	0	1	0	1	1	1	0	1	0
20	0	0	1	0	0	1	0	0	1
21	0	1	0	1	1	0	0	1	0
22	1	1	1	0	1	1	0	1	1
23	1	1	1	0	0	0	1	1	0
24	0	1	1	0	1	0	0	1	0
25	1	1	1	1	1	0	0	1	0
26	1	1	1	0	0	0	0	1	0
27	1	1	1	0	0	1	1	1	1
28	1	1	0	1	1	1	0	1	0
29	1	1	1	0	1	0	0	1	0
30	0	1	0	1	1	0	0	1	0

Tableau 4: Résultats du questionnaire concernant la Trace 3. Les techniques sont dénotées par A, M et V et les variables mesurées le sont par B, P et I.

DONNEES POUR LA TRACE 4									
Sujet #	A_B	A_P	A_I	M_B	M_P	M_I	V_B	V_P	V_I
1	0	0	0	0	1	0	1	1	1
2	0	1	1	1	1	1	1	1	1
3	0	1	0	1	1	1	1	1	1
4	0	1	0	1	1	0	1	1	1
5	0	1	0	1	1	0	1	1	1
6	0	0	0	1	1	0	1	1	1
7	0	0	0	1	1	0	1	1	1
8	0	0	0	1	1	1	1	1	1
9	0	1	0	1	1	1	1	1	1
10	0	1	0	1	1	0	1	1	1
11	0	0	0	1	1	1	1	1	1
12	1	1	1	1	1	1	0	1	1
13	0	1	0	1	1	1	1	1	1
14	0	1	0	1	1	1	1	1	1
15	0	1	0	0	0	0	1	1	0
16	0	1	0	0	1	0	1	1	1
17	1	0	1	1	1	1	1	1	1
18	1	0	1	1	1	1	1	1	1
19	0	0	0	1	1	1	1	1	0
20	0	0	0	0	0	1	1	0	1
21	0	0	1	1	1	1	1	1	1
22	0	0	0	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1
24	0	0	1	0	1	0	1	1	1
25	0	0	0	1	1	1	1	1	1
26	0	1	0	1	1	1	1	1	1
27	0	1	0	1	1	1	1	1	1
28	1	0	1	1	1	1	1	1	1
29	0	0	0	1	1	1	1	1	1
30	0	1	0	1	1	1	1	1	1

Tableau 5: Résultats du questionnaire concernant la Trace 4. Les techniques sont dénotées par A, M et V et les variables mesurées le sont par B, P et I.

DONNEES POUR LA TRACE 5									
Sujet #	A_B	A_P	A_I	M_B	M_P	M_I	V_B	V_P	V_I
1	1	1	1	0	0	1	1	1	1
2	0	1	1	0	1	1	0	1	1
3	0	1	1	0	1	1	0	1	1
4	0	1	1	1	1	1	0	1	1
5	0	1	1	1	1	1	1	1	1
6	0	1	1	1	1	0	1	1	1
7	1	1	1	1	1	1	1	1	1
8	1	1	1	0	1	1	1	0	1
9	1	1	1	1	1	1	1	1	1
10	1	1	0	0	1	1	1	1	1
11	0	1	1	0	1	1	0	1	1
12	1	1	1	1	1	1	0	1	1
13	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1
15	0	0	1	0	1	1	1	1	0
16	0	1	1	0	1	1	0	1	1
17	1	1	1	1	1	1	1	1	1
18	0	1	1	1	1	1	1	1	1
19	0	0	0	1	1	1	0	1	1
20	0	0	0	0	0	0	0	1	1
21	0	1	1	1	1	1	1	1	1
22	0	1	1	0	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	0	1	1
25	1	1	1	1	1	1	0	1	1
26	0	1	1	1	1	1	1	1	1
27	0	1	1	1	1	1	1	1	1
28	0	1	1	1	1	1	0	1	1
29	0	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	1

Tableau 6: Résultats du questionnaire concernant la Trace 5. Les techniques sont dénotées par A, M et V et les variables mesurées le sont par B, P et I.

DONNEES POUR LA TRACE 6									
Sujet #	A_B	A_P	A_I	M_B	M_P	M_I	V_B	V_P	V_I
1	0	0	0	1	0	1	1	1	1
2	0	0	1	1	1	1	1	1	1
3	0	0	0	1	0	1	1	1	1
4	0	0	0	1	1	1	1	1	1
5	0	0	0	0	0	0	1	1	1
6	0	0	0	0	0	0	0	1	1
7	1	0	0	1	1	1	1	1	1
8	0	0	0	1	1	1	1	1	1
9	0	0	1	1	0	1	1	1	1
10	0	0	0	0	0	0	1	1	1
11	0	0	0	0	0	1	1	1	1
12	1	1	1	1	1	1	1	1	1
13	0	0	0	1	1	1	1	1	1
14	0	1	1	1	1	1	1	1	1
15	0	0	0	0	0	0	1	1	1
16	0	0	1	0	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1
19	0	0	1	0	0	1	1	1	0
20	0	0	0	0	0	1	1	1	1
21	0	1	1	1	1	1	1	1	1
22	0	0	1	1	1	1	1	1	1
23	0	1	1	1	1	1	1	1	1
24	0	1	0	1	0	1	1	1	1
25	0	0	1	1	1	1	1	1	1
26	0	0	0	1	1	1	1	1	1
27	1	0	1	1	1	1	1	1	1
28	0	0	0	1	1	1	1	1	1
29	0	1	1	1	1	1	1	1	1
30	0	0	1	1	1	1	1	1	1

Tableau 7: Résultats du questionnaire concernant la Trace 6. Les techniques sont dénotées par A, M et V et les variables mesurées le sont par B, P et I.

Annexe E

Programmes informatiques

Les prochaines pages présentent les programmes qui ont été écrits afin d'effectuer les différentes analyses statistiques.

Premièrement, un premier programme a été développé en Borland PASCAL (version 7.0) sur un micro-ordinateur Pentium de DELL afin de faciliter les calculs reliés au développement d'un test exact associé avec la méthode du maximum de vraisemblance discuté à la section 5.2. Il sert à calculer toutes les permutations possibles de la statistique T et les probabilités reliées à chacune d'elle. De cette façon, la distribution conditionnelle exacte de T sous H_0 a pu être obtenue.

Ensuite, les différents programmes développés en SAS sont exposés. Ces programmes ont été utilisés pour faire l'analyse des données avec le test de Student (test t) et la régression logistique.

Finalement, les logiciels MATHEMATICA et RESAMPLING STATS ont aidé à effectuer certaines manipulations reliées au ré-échantillonnage. Le logiciel MATHEMATICA a permis de faire les calculs pour la création des intervalles de confiance de Clopper-Pearson. Le logiciel RESAMPLING STATS a pour sa part été très utile pour la génération de milliers de nouveaux échantillons *bootstrap* obtenus à partir des données de l'échantillon initial.

Le programme suivant a été développé en Borland PASCAL. Il a permis de déterminer la distribution de la statistique du maximum de vraisemblance et la *p-value* associée à chaque comparaison simple pour tous les tracés de l'étude.

```
PROGRAM Permutations;

VAR
  FichierOUT :      TEXT;

FUNCTION Facto(N : INTEGER) : REAL;

VAR
  Count :      INTEGER;
  X :        REAL;

BEGIN
  X := 1;
  FOR Count := 1 to N DO
    BEGIN
      X := X*Count;
    END;
  Facto := X;
END;

FUNCTION Comb(c : INTEGER; N : INTEGER) : REAL;

BEGIN
  IF (c <> 0) AND (N <> 0)
  THEN
    Comb := Facto(N)/(Facto(c)*Facto(N-c))
  ELSE
    Comb := 1;
  END;
END;

FUNCTION Power(base : REAL;N : INTEGER) : REAL;

VAR
  X :        REAL;
  Count :    INTEGER;

BEGIN
  X := 1;
  FOR Count := 1 TO N DO
    X := X * base;
  Power := X;
END;
```

```

PROCEDURE CalculProbabilite;

VAR
  Count1,Count2 :      INTEGER;
  Prob,T :           REAL;
  N :               INTEGER;

BEGIN
  FOR N := 0 TO 30 DO
    BEGIN
      FOR Count1 := 0 TO N DO
        BEGIN
          {Calcul la probabilite}
          Count2:= N-Count1;
          Prob := Comb(Count1,N)*Comb(Count2,Count2);
          Prob := Prob/Power(2,N);
          WRITELN(FichierOUT,N:4,Count1:4,Count2:4,' ',Prob:12:10);
        END;
      END;
    END;
  END;

PROCEDURE CalculP_Value;

VAR
  FichierIN1 : TEXT;
  FichierIN2 : TEXT;
  FichierOUT : TEXT;
  Ntotal1,Nci1 : INTEGER;
  Ntotal2,Nci2,Nic2 : INTEGER;
  Prob2 : REAL;
  Prob : ARRAY[0..30] OF REAL;
  P_value : REAL;
  Count1,Count2 : INTEGER;

BEGIN
  ASSIGN(FichierIN1,'C:\CARO\PROGRAMS\DATA.TXT');
  RESET(FichierIN1);
  ASSIGN(FichierIN2,'C:\CARO\PROGRAMS\PROB.TXT');
  ASSIGN(FichierOUT,'C:\CARO\PROGRAMS\P_VALUE.TXT');
  REWRITE(FichierOUT);
  FOR Count1 := 1 TO 54 DO
    BEGIN
      RESET(FichierIN2);
      READLN(FichierIN1,Ntotal1,Nci1);
      READLN(FichierIN2,Ntotal2);

      WHILE (Ntotal1 <> Ntotal2) DO
        BEGIN
          READLN(FichierIN2,Ntotal2,Nci2,Nic2,Prob2);
        END;
    END;
  END;

```

```

FOR Count2 := 0 TO 30 DO
  BEGIN
    Prob[Count2] := 0;
  END;
FOR Count2 := 0 TO Ntotal1 DO
  BEGIN
    Prob[Count2] := Prob2;
    READLN(FichierIN2,Ntotal2,Nci2,Nic2,Prob2);
  END;

P_Value := 0;
FOR Count2:= 0 TO Nci1 DO
  BEGIN
    P_Value:= P_Value + Prob[Count2];
  END;

  WRITELN(FichierOUT,Ntotal1:4,Nci1:4,' ',P_value:10:8);
END;
CLOSE(FichierIN1);
CLOSE(FichierIN2);
CLOSE(FichierOUT);
END;

BEGIN
  ASSIGN(FichierOUT,'C:\CARO\PROGRAMS\PROB.TXT');
  REWRITE(FichierOUT);
  CalculProbabilite;
  CLOSE(FichierOUT);
  CalculP_Value;
END.

```

Le programme suivant, développé en SAS, a permis de calculer la statistique t pour chaque variable et pour chacun des 6 tracés de l'étude.

```
/* *****  
/* TRACE1 Trace1 TRACE1 Trace1 TRACE1 Trace1 TRACE1 Trace1 TRACE1 Trace1 */  
/* *****  
LIBNAME Trace1 '/home/anctil/These/Trace1';  
  
DATA Trace1;  
  SET Trace1.Combine1;  
  DiffBam = Ba-Bm;  
  DiffBmv = Bm-Bv;  
  DiffPam = Pa-Pm;  
  DiffPmv = Pm-Pv;  
  DiffIam = Ia-Bm;  
  DiffImv = Im-Bv;  
  OUTPUT;  
  
PROC MEANS DATA=Trace1 N MEAN STDERR STD MIN MAX T PRT;  
  VAR DIFFBam DIFFBmv DIFFPam DIFFPmv DIFFIam DIFFImv;  
  TITLE 'Paired comparisons for TRACE1';  
  
run;  
quit;
```

Le programme suivant, développé en SAS, a permis de réaliser la régression logistique. Ce programme a été exécuté pour chacun des 6 tracés de l'étude.

```

/* ***** */
/* Comparaison de A versus M */
/* Yam = 0 si (A=1 et M=1) ou (A=0 et M=0)
   Yam = 1 si (A=0 et M=1)
   Yam = 2 si (A=1 et M=0) */

/* Comparaison de M versus V */
/* Ymv = 0 si (M=1 et V=1) ou (A=0 et V=0)
   Ymv = 1 si (M=0 et V=1)
   Ymv = 2 si (M=1 et V=0) */

/* Comparaison de A versus M versus V */
/* Yamv = 0 si (A=1 et M=1 et V=1) ou (A=0 et M=0 et V=0)
   Yamv = 1 si (A=1 et M=1 et V=0)
   Yamv = 2 si (A=1 et M=0 et V=1)
   Yamv = 3 si (A=1 et M=0 et V=0)
   Yamv = 4 si (A=0 et M=1 et V=1)
   Yamv = 5 si (A=0 et M=0 et V=1)
   Yamv = 6 si (A=0 et M=0 et V=0) */
/* ***** */

/* ***** */
/* TRACE1 Trace1 TRACE1 Trace1 TRACE1 Trace1 TRACE1 Trace1 Trace1 Trace1 */
/* ***** */
LIBNAME Trace1 '/home/anctil/These/Trace1';

DATA Combine1;
  SET Trace1.Combine1;
  OUTPUT;

/* Comparaisons A vs M ---- Variable B*/
DATA Combine1;
  SET Combine1;
  IF Ba=1 AND Bm=1 THEN YBam=0;
  IF Ba=0 AND Bm=0 THEN YBam=0;
  IF Ba=0 AND Bm=1 THEN YBam=1;
  IF Ba=1 AND Bm=0 THEN YBam=2;

/* Comparaisons A vs M ---- Variable P*/

  IF Pa=1 AND Pm=1 THEN YPam=0;
  IF Pa=0 AND Pm=0 THEN YPam=0;
  IF Pa=0 AND Pm=1 THEN YPam=1;
  IF Pa=1 AND Pm=0 THEN YPam=2;

/* Comparaisons A vs M ---- Variable I*/

```

```

IF Ia=1 AND Im=1 THEN YIam=0;
IF Ia=0 AND Im=0 THEN YIar=0;
IF Ia=0 AND Im=1 THEN YIam=1;
IF Ia=1 AND Im=0 THEN YIam=2;

/* Comparaisons M vs V ---- Variable B*/

IF Bm=1 AND Bv=1 THEN YBmv=0;
IF Bm=0 AND Bv=0 THEN YBmv=0;
IF Bm=0 AND Bv=1 THEN YBmv=1;
IF Bm=1 AND Bv=0 THEN YBmv=2;

/* Comparaisons M vs V ---- Variable P*/

IF Pm=1 AND Pv=1 THEN YPmv=0;
IF Pm=0 AND Pv=0 THEN YPmv=0;
IF Pm=0 AND Pv=1 THEN YPmv=1;
IF Pm=1 AND Pv=0 THEN YPmv=2;

/* Comparaisons M vs V ---- Variable I*/

IF Im=1 AND Iv=1 THEN YImv=0;
IF Im=0 AND Iv=0 THEN YImv=0;
IF Im=0 AND Iv=1 THEN YImv=1;
IF Im=1 AND Iv=0 THEN YImv=2;

/* Comparaisons A vs M vs V --- Variable B */

IF Ba=1 AND Bm=1 AND Bv=1 THEN YBamv=0;
IF Ba=0 AND Bm=0 AND Bv=0 THEN YBamv=0;
IF Ba=1 AND Bm=1 AND Bv=0 THEN YBamv=1;
IF Ba=1 AND Bm=0 AND Bv=1 THEN YBamv=2;
IF Ba=1 AND Bm=0 AND Bv=0 THEN YBamv=3;
IF Ba=0 AND Bm=1 AND Bv=1 THEN YBamv=4;
IF Ba=0 AND Bm=1 AND Bv=0 THEN YBamv=5;
IF Ba=0 AND Bm=0 AND Bv=1 THEN YBamv=6;

/* Comparaisons A vs M vs V --- Variable P */

IF Pa=1 AND Pm=1 AND Pv=1 THEN YPamv=0;
IF Pa=0 AND Pm=0 AND Pv=0 THEN YPamv=0;
IF Pa=1 AND Pm=1 AND Pv=0 THEN YPamv=1;
IF Pa=1 AND Pm=0 AND Pv=1 THEN YPamv=2;
IF Pa=1 AND Pm=0 AND Pv=0 THEN YPamv=3;
IF Pa=0 AND Pm=1 AND Pv=1 THEN YPamv=4;
IF Pa=0 AND Pm=1 AND Pv=0 THEN YPamv=5;
IF Pa=0 AND Pm=0 AND Pv=1 THEN YPamv=6;

/* Comparaisons A vs M vs V --- Variable I */

```

```

IF Ia=1 AND Im=1 AND Iv=1 THEN YIamv=0;
IF Ia=0 AND Im=0 AND Iv=0 THEN YIamv=0;
IF Ia=1 AND Im=1 AND Iv=0 THEN YIamv=1;
IF Ia=1 AND Im=0 AND Iv=1 THEN YIamv=2;
IF Ia=1 AND Im=0 AND Iv=0 THEN YIamv=3;
IF Ia=0 AND Im=1 AND Iv=1 THEN YIamv=4;
IF Ia=0 AND Im=1 AND Iv=0 THEN YIamv=5;
IF Ia=0 AND Im=0 AND Iv=1 THEN YIamv=6;

DATA Trace1.Log1;
  SET Combine1;

LIBNAME Trace  '/home/anctil/These/CorrectLog';

DATA YBam1;
  SET Trace.YBam1;
  wwweight=1; IF YBam=0 THEN wwweight=0.000000000001;
  OUTPUT;

DATA YPmv1;
  SET Trace.YPmv1;
  wwweight=1; IF YPmv=0 THEN wwweight=0.000000000001;
  OUTPUT;

DATA YImv1;
  SET Trace.YImv1;
  wwweight=1; IF YImv=0 THEN wwweight=0.000000000001;
  OUTPUT;

DATA YBamv1;
  SET Trace.YBamv1;
  wwweight=1; IF YBamv=1 | YBamv=2 THEN wwweight=0.00000001;
  IF YBamv=1 THEN YBamv=10;
  IF YBamv=2 THEN YBamv=11;
  OUTPUT;

DATA YPamv1;
  SET Trace.YPamv1;
  wwweight=1; IF YPamv=0 | YPamv=2 | YPamv=4 THEN wwweight=0.00000001;
  IF YPamv=0 THEN YPamv=10;
  IF YPamv=2 THEN YPamv=11;
  IF YPamv=4 THEN YPamv=12;
  OUTPUT;

DATA YIamv1;
  SET Trace.YIamv1;
  wwweight=1; IF YIamv=0 | YIamv=4 THEN wwweight=0.00000001;
  IF YIamv=0 THEN YIamv=10;
  IF YPamv=4 THEN YPamv=11;
  OUTPUT;

```

```

DATA Log1;
  SET Trace.Log1;

/***** VARIABLE B *****/

PROC LOGISTIC DATA=YBam1;
  MODEL YBam= /RiskLimits;
  Weight wwweight;

PROC LOGISTIC DATA=Log1F;
  MODEL YBmv= /RiskLimits;

PROC LOGISTIC DATA=YBamv1;
  MODEL YBamv= /RiskLimits;
  Weight wwweight;

/***** VARIABLE P *****/

PROC LOGISTIC DATA=Log1F;
  MODEL YPam= /RiskLimits;

PROC LOGISTIC DATA=YPmv1;
  MODEL YPmv= /RiskLimits;
  Weight wwweight;

PROC LOGISTIC DATA=YPamv1;
  MODEL YPamv= /RiskLimits;
  Weight wwweight;

/***** VARIABLE I *****/

PROC LOGISTIC DATA=Log1F;
  MODEL YIam= /RiskLimits;

PROC LOGISTIC DATA=YImv1;
  MODEL YImv= /RiskLimits;
  Weight wwweight;

PROC LOGISTIC DATA=YIamv1;
  MODEL YIamv= /RiskLimits;
  Weight wwweight;

run;
quit;

```

Le programme suivant a servi à générer les bornes inférieures et supérieures de Clopper-Pearson à un niveau $\alpha = 0.05$. Le logiciel MATHEMATICA a exécuté cette tâche.

```

Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,0}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,1}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,2}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,3}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,4}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,5}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,6}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,7}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,8}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,9}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,10}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,11}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,12}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,13}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,14}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,15}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,16}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,17}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,18}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,19}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,20}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,0,21}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),

```



```

Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,19,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,20,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,21,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,22,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,23,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,24,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,25,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,26,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,27,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,28,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,29,30}] == 0.025, x]
Solve[Sum[Factorial[30]/(Factorial[i]*Factorial[30-i])x^(i)*(1-x)^(30-i),
{i,30,30}] == 0.025, x]

```

Le logiciel RESAMPLING STATS a permis de faire le ré-échantillonnage *bootstrap* relié à l'étude. Le programme suivant donne les étapes spécifiques de cette méthode statistique.

```
'Simulating Clopper-Pearson intervals for Westfall-Young resampling  
'procedure --- book pg 177-180
```

```
COPY (5000) NUM 'number of resamples to do (not permutations!
```

```
'set up which y lead to CP intervals containing the observed values
```

```
'input those observable y whose CP interval  
'contains the observed success proportion for A  
COPY (3 4 5 6 7 8 9 10 11 12) AYESVECT
```

```
COPY (19 20 21 22 23 24 25 26 27 28) MYESVECT  
COPY (25 26 27 28 29 30) VYESVECT
```

```
'copy (a m v) r  
COPY (0 1 1) R1  
COPY (0 1 1) R2  
COPY (0 1 0) R3  
COPY (0 1 1) R4  
COPY (0 1 1) R5  
COPY (0 0 1) R6  
COPY (0 1 1) R7  
COPY (0 1 1) R8  
COPY (1 1 1) R9  
COPY (1 1 0) R10  
COPY (0 1 1) R11  
COPY (0 1 1) R12  
COPY (1 1 1) R13  
COPY (0 1 1) R14  
COPY (0 0 1) R15  
COPY (0 0 1) R16  
COPY (0 1 1) R17  
COPY (0 1 1) R18  
COPY (0 1 1) R19  
COPY (0 1 1) R20  
COPY (0 1 1) R21  
COPY (0 0 1) R22  
COPY (1 1 1) R23  
COPY (0 0 1) R24  
COPY (1 1 1) R25  
COPY (1 1 1) R26  
COPY (0 0 1) R27  
COPY (1 1 1) R28  
COPY (0 1 1) R29  
COPY (0 1 1) R30
```

REPEAT NUM
SHUFFLE R1 S1
SHUFFLE R2 S2
SHUFFLE R3 S3
SHUFFLE R4 S4
SHUFFLE R5 S5
SHUFFLE R6 S6
SHUFFLE R7 S7
SHUFFLE R8 S8
SHUFFLE R9 S9
SHUFFLE R10 S10
SHUFFLE R11 S11
SHUFFLE R12 S12
SHUFFLE R13 S13
SHUFFLE R14 S14
SHUFFLE R15 S15
SHUFFLE R16 S16
SHUFFLE R17 S17
SHUFFLE R18 S18
SHUFFLE R19 S19
SHUFFLE R20 S20
SHUFFLE R21 S21
SHUFFLE R22 S22
SHUFFLE R23 S23
SHUFFLE R24 S24
SHUFFLE R25 S25
SHUFFLE R26 S26
SHUFFLE R27 S27
SHUFFLE R28 S28
SHUFFLE R29 S29
SHUFFLE R30 S30

TAKE S1 1 A1
TAKE S1 2 M1
TAKE S1 3 V1
TAKE S2 1 A2
TAKE S2 2 M2
TAKE S2 3 V2
TAKE S3 1 A3
TAKE S3 2 M3
TAKE S3 3 V3
TAKE S4 1 A4
TAKE S4 2 M4
TAKE S4 3 V4
TAKE S5 1 A5
TAKE S5 2 M5
TAKE S5 3 V5
TAKE S6 1 A6
TAKE S6 2 M6
TAKE S6 3 V6

TAKE S7 1 A7
 TAKE S7 2 M7
 TAKE S7 3 V7
TAKE S8 1 A8
 TAKE S8 2 M8
 TAKE S8 3 V8
TAKE S9 1 A9
 TAKE S9 2 M9
 TAKE S9 3 V9
TAKE S10 1 A10
 TAKE S10 2 M10
 TAKE S10 3 V10
TAKE S11 1 A11
 TAKE S11 2 M11
 TAKE S11 3 V11
TAKE S12 1 A12
 TAKE S12 2 M12
 TAKE S12 3 V12
TAKE S13 1 A13
 TAKE S13 2 M13
 TAKE S13 3 V13
TAKE S14 1 A14
 TAKE S14 2 M14
 TAKE S14 3 V14
TAKE S15 1 A15
 TAKE S15 2 M15
 TAKE S15 3 V15
TAKE S16 1 A16
 TAKE S16 2 M16
 TAKE S16 3 V16
TAKE S17 1 A17
 TAKE S17 2 M17
 TAKE S17 3 V17
TAKE S18 1 A18
 TAKE S18 2 M18
 TAKE S18 3 V18
TAKE S19 1 A19
 TAKE S19 2 M19
 TAKE S19 3 V19
TAKE S20 1 A20
 TAKE S20 2 M20
 TAKE S20 3 V20
TAKE S21 1 A21
 TAKE S21 2 M21
 TAKE S21 3 V21
TAKE S22 1 A22
 TAKE S22 2 M22
 TAKE S22 3 V22
TAKE S23 1 A23
 TAKE S23 2 M23
 TAKE S23 3 V23

```

TAKE S24 1 A24
      TAKE S24 2 M24
            TAKE S24 3 V24
TAKE S25 1 A25
      TAKE S25 2 M25
            TAKE S25 3 V25
TAKE S26 1 A26
      TAKE S26 2 M26
            TAKE S26 3 V26
TAKE S27 1 A27
      TAKE S27 2 M27
            TAKE S27 3 V27
TAKE S28 1 A28
      TAKE S28 2 M28
            TAKE S28 3 V28
TAKE S29 1 A29
      TAKE S29 2 M29
            TAKE S29 3 V29
TAKE S30 1 A30
      TAKE S30 2 M30
            TAKE S30 3 V30

```

```

CONCAT A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15 A16 A17
      A18 A19 A20 A21 A22 A23 A24 A25 A26 A27 A28 A29 A30 A
COUNT A =1 AOUT
COUNT AOUT MEMBEROF AYESVECT AYES 'did the resampled CP int contain

```

```

CONCAT M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 M13 M14 M15 M16 M17
      M18 M19 M20 M21 M22 M23 M24 M25 M26 M27 M28 M29 M30 M
COUNT M =1 MOUT
COUNT MOUT MEMBEROF MYESVECT MYES

```

```

CONCAT V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17
      V18 V19 V20 V21 V22 V23 V24 V25 V26 V27 V28 V29 V30 V
COUNT V =1 VOUT
COUNT VOUT MEMBEROF VYESVECT VYES

```

```

MULTIPLY AYES MYES HIT1
MULTIPLY MYES VYES HIT2
MULTIPLY AYES MYES VYES HIT

```

```

SCORE HIT1 TOT1
SCORE HIT2 TOT2
SCORE HIT TOT

```

```

END
HISTOGRAM TOT1
COUNT TOT1 =1 KK1
DIVIDE KK1 NUM PVALUE1
PRINT PVALUE1

```

```
HISTOGRAM TOT2  
COUNT TOT2 =1 KK2  
DIVIDE KK2 NUM PVALUE2  
PRINT PVALUE2
```

```
HISTOGRAM TOT  
COUNT TOT =1 KK  
DIVIDE KK NUM PVALUE  
PRINT PVALUE
```

Annexe F

Table de contingence

		Question 2		
		correcte	incorrecte	
Question 1	correcte	n_{cc}	n_{ci}	$n_{c.}$
	incorrecte	n_{ic}	n_{ii}	$n_{i.}$
		$n_{.c}$	$n_{.i}$	n

Tableau 8: Table de contingence utilisée lors de la comparaison de 2 variables (questions) pour un même tracé. Les termes n_{xy} ($x, y \in \{c, i\}$) donnent le nombre de personnes ayant un résultat x à la première question **ET** un résultat y à la deuxième question. L'indice x représente le résultat de la première question et l'indice y représente le résultat de la deuxième question. Les indices x et y ont la valeur "c" si la réponse est "correcte" et "i" si la réponse est "incorrecte".

Annexe G

Estimateurs du maximum de vraisemblance calculés avec MATHEMATICA

Ce qui suit présente quelques résultats de la résolution d'équations effectuée par le logiciel MATHEMATICA afin de calculer les estimateurs de vraisemblance sous $H_0 : p_{c..} = p_{.c.} = p_{..c}$.

Afin de simplifier la notation, les paramètres ont été renommés comme suit,

$$p_{ccc} = A \text{ et } n_{ccc} = a$$

$$p_{cic} = B \text{ et } n_{cic} = b$$

$$p_{cci} = C \text{ et } n_{cci} = c$$

$$p_{cii} = D \text{ et } n_{cii} = d$$

$$p_{icc} = E \text{ et } n_{icc} = e$$

$$p_{ici} = F \text{ et } n_{ici} = f$$

$$p_{iic} = G \text{ et } n_{iic} = g$$

$$p_{iii} = H \text{ et } n_{iii} = h$$

La commande suivante, donnée à MATHEMATICA, vise à résoudre la première dérivée du log de vraisemblance par rapport à p_{ccc} (A).

$$\text{Solve}[a/A - h/(1-A-3E-2F+D-2G) == 0, A]$$

La solution est alors

$$A \rightarrow -\frac{-a-aD+3ae+2aF+2aG}{a+h}$$

Une autre commande permet de résoudre la première dérivée du log de vraisemblance par rapport à p_{ici} (F).

$$\text{Solve}[b/(E + F -D) + f/F - 2h/(1-A-3E-2F+D-2G) == 0, F]$$

Les solutions sont

$$F \rightarrow \frac{Z_1 \pm \sqrt{-4f(D-AD+D^2-c+Ac-ADc+3e^2-2DG+2cG)(Z_2)+(-Z_1)^2}}{2(Z_2)}$$

où

$$Z_1 = b - Ab + bD - 3be + f - Af + 3Df - 5ef - 2bG - 2fG + 2Dh - 2ch$$

$$Z_2 = 2b + 2f + 2h$$

Il est évident que cette solution ne permettrait pas de créer un test simple basé sur la méthode du maximum de vraisemblance pour vérifier des hypothèses multiples. Donc, cette option a été abandonnée.

Annexe H

Intervalles de confiance de Clopper-Pearson

Le tableau de la page suivante contient les intervalles de confiance de Clopper-Pearson pour $b_j = 0, \dots, 30$ à un niveau $\alpha = 0.05$. Ces intervalles sont définis comme suit :

$$I_\alpha(b_j) = \begin{cases} [0, \pi_S(b_j)), & \text{pour } b_j = 0 \\ (\pi_I(b_j), \pi_S(b_j)), & \text{pour } 0 < b_j < n \\ (\pi_I(b_j), 1], & \text{pour } b_j = n \end{cases}$$

$I_{0.05}(b_j = i)$	Borne Inférieure $\pi_I(b_j)$	Borne Supérieure $\pi_S(b_j)$
$I_{0.05}(b_j = 0)$	[0	0.11570)
$I_{0.05}(b_j = 1)$	(0.00084	0.17217)
$I_{0.05}(b_j = 2)$	(0.00817	0.22074)
$I_{0.05}(b_j = 3)$	(0.02112	0.26529)
$I_{0.05}(b_j = 4)$	(0.03755	0.30722)
$I_{0.05}(b_j = 5)$	(0.05642	0.34721)
$I_{0.05}(b_j = 6)$	(0.07714	0.38567)
$I_{0.05}(b_j = 7)$	(0.09934	0.42284)
$I_{0.05}(b_j = 8)$	(0.12280	0.45889)
$I_{0.05}(b_j = 9)$	(0.14735	0.49396)
$I_{0.05}(b_j = 10)$	(0.17287	0.52812)
$I_{0.05}(b_j = 11)$	(0.19930	0.56144)
$I_{0.05}(b_j = 12)$	(0.22656	0.59397)
$I_{0.05}(b_j = 13)$	(0.25461	0.62573)
$I_{0.05}(b_j = 14)$	(0.28342	0.65674)
$I_{0.05}(b_j = 15)$	(0.31297	0.68703)
$I_{0.05}(b_j = 16)$	(0.34326	0.71658)
$I_{0.05}(b_j = 17)$	(0.37427	0.74539)
$I_{0.05}(b_j = 18)$	(0.40604	0.77344)
$I_{0.05}(b_j = 19)$	(0.43856	0.80070)
$I_{0.05}(b_j = 20)$	(0.47188	0.82713)
$I_{0.05}(b_j = 21)$	(0.50604	0.85265)
$I_{0.05}(b_j = 22)$	(0.54111	0.87721)
$I_{0.05}(b_j = 23)$	(0.57716	0.90066)
$I_{0.05}(b_j = 24)$	(0.61433	0.92286)
$I_{0.05}(b_j = 25)$	(0.65279	0.94358)
$I_{0.05}(b_j = 26)$	(0.69278	0.96245)
$I_{0.05}(b_j = 27)$	(0.73471	0.97888)
$I_{0.05}(b_j = 28)$	(0.77927	0.99182)
$I_{0.05}(b_j = 29)$	(0.82783	0.99916)
$I_{0.05}(b_j = 30)$	(0.88430	1]

Tableau 9: Intervalles de confiance de Clopper-Pearson pour $\alpha = 0.05$.

Annexe I

Présentation des résultats

Les pages qui suivent présentent un résumé des résultats (*p-values*) des analyses statistiques pour les 6 tracés employés et ce, pour les méthodes statistiques applicables selon le cas. Les 3 prochaines sections contiennent l'information suivante,

- Section 1. Analyse univariée (comparaisons simples)
- Section 2. Analyse univariée (comparaisons multiples)
- Section 3. Résultats du test de McNemar

I.1 Analyse univariée (Comparaisons simples)

Les pages qui suivent présentent un résumé des résultats des comparaisons simples appliquées dans le cas d'une analyse univariée.

SOMMAIRE DES RESULTATS POUR LA TRACE 1				
Variable	Technique	A vs M		M vs V
<i>B</i>	Test <i>t</i>	0.00005	$p(A) < p(M)$	0.0804 $p(M) = p(V)$
	Max. Vrais.	0.00001	$p(A) < p(M)$	0.1445 $p(M) = p(V)$
	Ré-échantill.	0.0006	$p(A) \neq p(M)$	0.0030 $p(M) \neq p(V)$
	Régr. Log.	1.0000	$p(A) = p(M)$	0.0027 $p(M) \neq p(V)$
<i>P</i>	Test <i>t</i>	0.0028	$p(A) < p(M)$	0.0030 $p(M) < p(V)$
	Max. Vrais.	0.0064	$p(A) < p(M)$	0.0078 $p(M) < p(V)$
	Ré-échantill.	0.0304	$p(A) \neq p(M)$	0.0020 $p(M) \neq p(V)$
	Régr. Log.	0.00007	$p(A) \neq p(M)$	1.0000 $p(M) = p(V)$
<i>I</i>	Test <i>t</i>	0.0001	$p(A) < p(M)$	0.00005 $p(M) < p(V)$
	Max. Vrais.	0.5000	$p(A) = p(M)$	0.00003 $p(M) < p(V)$
	Ré-échantill.	0.2344	$p(A) = p(M)$	0.0050 $p(M) \neq p(V)$
	Régr. Log.	0.00001	$p(A) \neq p(M)$	1.0000 $p(M) = p(V)$

Tableau 10: Sommaire des *p-values* pour la Trace 1 selon les différentes techniques statistiques utilisées pour chaque comparaison simple.

SOMMAIRE DES RESULTATS POUR LA TRACE 2					
Variable	Technique	A vs M		M vs V	
<i>B</i>	Test <i>t</i>	0.00005	$p(A) < p(M)$	0.8372	$p(M) = p(V)$
	Max. de Vrais.	0.00002	$p(A) < p(M)$	0.9102	$p(M) = p(V)$
	Ré-échantill.	0.0224	$p(A) \neq p(M)$	0.1344	$p(M) = p(V)$
	Régr. Log.	1.0000	$p(A) = p(M)$	0.0007	$p(M) \neq p(V)$
<i>P</i>	Test <i>t</i>	0.0150	$p(A) < p(M)$	0.0085	$p(M) < p(V)$
	Max. de Vrais.	0.0287	$p(A) < p(M)$	0.0195	$p(M) < p(V)$
	Ré-échantill.	0.1580	$p(A) = p(M)$	0.0830	$p(M) = p(V)$
	Régr. Log.	0.0001	$p(A) \neq p(M)$	0.0122	$p(M) \neq p(V)$
<i>I</i>	Test <i>t</i>	0.0162	$p(A) < p(M)$	0.3934	$p(M) = p(V)$
	Max. de Vrais.	0.2744	$p(A) = p(M)$	0.6563	$p(M) = p(V)$
	Ré-échantill.	0.9292	$p(A) = p(M)$	0.9894	$p(M) = p(V)$
	Régr. Log.	0.0002	$p(A) \neq p(M)$	0.0039	$p(M) \neq p(V)$

Tableau 11: Sommaire des *p-values* pour la Trace 2 selon les différentes techniques statistiques utilisées pour chaque comparaison simple.

SOMMAIRE DES RESULTATS POUR LA TRACE 3					
Variable	Technique	A vs M		M vs V	
<i>B</i>	Test <i>t</i>	0.5000	$p(A) = p(M)$	0.9913	$p(M) = p(V)$
	Max. Vrais.	0.6047	$p(A) = p(M)$	0.9963	$p(M) = p(V)$
	Ré-échantill.	0.7710	$p(A) = p(M)$	0.3950	$p(M) = p(V)$
	Régr. Log.	0.00001	$p(A) \neq p(M)$	0.0001	$p(M) \neq p(V)$
<i>P</i>	Test <i>t</i>	0.9080	$p(A) = p(M)$	0.0113	$p(M) < p(V)$
	Max. Vrais.	0.9688	$p(A) = p(M)$	0.0313	$p(M) < p(V)$
	Ré-échantill.	0.9572	$p(A) = p(M)$	0.5626	$p(M) = p(V)$
	Régr. Log.	0.0253	$p(A) \neq p(M)$	1.0000	$p(M) = p(V)$
<i>I</i>	Test <i>t</i>	0.5935	$p(A) = p(M)$	0.9783	$p(M) = p(V)$
	Max. Vrais.	0.7728	$p(A) = p(M)$	0.9990	$p(M) = p(V)$
	Ré-échantill.	0.8448	$p(A) = p(M)$	0.4124	$p(M) = p(V)$
	Régr. Log.	0.0000	$p(A) \neq p(M)$	0.0108	$p(M) \neq p(V)$

Tableau 12: Sommaire des *p-values* pour la Trace 3 selon les différentes techniques statistiques utilisées pour chaque comparaison simple.

SOMMAIRE DES RESULTATS POUR LA TRACE 4					
Variable	Technique	A vs M		M vs V	
<i>B</i>	Test <i>t</i>	0.00005	$p(A) < p(M)$	0.0517	$p(M) = p(V)$
	Max. Vrais.	0.0000	$p(A) < p(M)$	0.1094	$p(M) = p(V)$
	Ré-échantill.	0.0000	$p(A) \neq p(M)$	0.0000	$p(M) \neq p(V)$
	Régr. Log.	1.0000	$p(A) = p(M)$	0.0201	$p(M) \neq p(V)$
<i>P</i>	Test <i>t</i>	0.0001	$p(A) < p(M)$	0.1628	$p(M) = p(V)$
	Max. Vrais.	0.0005	$p(A) < p(M)$	0.5000	$p(M) = p(V)$
	Ré-échantill.	0.0270	$p(A) \neq p(M)$	0.0088	$p(M) \neq p(V)$
	Régr. Log.	0.0067	$p(A) \neq p(M)$	1.0000	$p(M) = p(V)$
<i>I</i>	Test <i>t</i>	0.00005	$p(A) < p(M)$	0.0045	$p(M) < p(V)$
	Max. Vrais.	0.0005	$p(A) < p(M)$	0.0195	$p(M) < p(V)$
	Ré-échantill.	0.0058	$p(A) \neq p(M)$	0.0034	$p(M) \neq p(V)$
	Régr. Log.	0.0067	$p(A) \neq p(M)$	0.0122	$p(M) \neq p(V)$

Tableau 13: Sommaire des *p-value* pour la Trace 4 selon les différentes techniques statistiques utilisées pour chaque comparaison simple.

SOMMAIRE DES RESULTATS POUR LA TRACE 5					
Variable	Technique	A vs M		M vs V	
<i>B</i>	Test <i>t</i>	0.0252	$p(A) < p(M)$	0.6157	$p(M) = p(V)$
	Max. Vrais.	0.0461	$p(A) < p(M)$	0.7256	$p(M) = p(V)$
	Ré-échantill.	0.6838	$p(A) = p(M)$	0.8910	$p(M) = p(V)$
	Régr. Log.	0.0002	$p(A) \neq p(M)$	0.0001	$p(M) \neq p(V)$
<i>P</i>	Test <i>t</i>	0.2863	$p(A) = p(M)$	0.2863	$p(M) = p(V)$
	Max. Vrais.	0.5000	$p(A) = p(M)$	0.5000	$p(M) = p(V)$
	Ré-échantill.	1.0000	$p(A) = p(M)$	0.9222	$p(M) = p(V)$
	Régr. Log.	0.0507	$p(A) = p(M)$	0.0507	$p(M) = p(V)$
<i>I</i>	Test <i>t</i>	0.9916	$p(A) = p(M)$	0.9977	$p(M) = p(V)$
	Max. Vrais.	0.5000	$p(A) = p(M)$	0.5000	$p(M) = p(V)$
	Ré-échantill.	0.7746	$p(A) = p(M)$	0.9190	$p(M) = p(V)$
	Régr. Log.	0.0507	$p(A) = p(M)$	0.0507	$p(M) = p(V)$

Tableau 14: Sommaire des *p-values* pour la Trace 5 selon les différentes techniques statistiques utilisées pour chaque comparaison simple.

SOMMAIRE DES RESULTATS POUR LA TRACE 6					
Variable	Technique	A vs M		M vs V	
<i>B</i>	Test <i>t</i>	0.00005	$p(A) < p(M)$	0.0030	$p(M) < p(V)$
	Max. Vrais.	0.00001	$p(A) < p(M)$	0.0078	$p(M) < p(V)$
	Ré-échantill.	0.0000	$p(A) \neq p(M)$	0.0002	$p(M) \neq p(V)$
	Régr. Log.	1.0000	$p(A) = p(M)$	1.0000	$p(M) = p(V)$
<i>P</i>	Test <i>t</i>	0.0006	$p(A) < p(M)$	0.0002	$p(M) < p(V)$
	Max. Vrais.	0.0017	$p(A) < p(M)$	0.0005	$p(M) < p(V)$
	Ré-échantill.	0.0050	$p(A) \neq p(M)$	0.0000	$p(M) \neq p(V)$
	Régr. Log.	0.0079	$p(A) \neq p(M)$	1.0000	$p(M) = p(V)$
<i>I</i>	Test <i>t</i>	0.0162	$p(A) < p(M)$	0.0416	$p(M) < p(V)$
	Max. Vrais.	0.0005	$p(A) < p(M)$	0.1875	$p(M) = p(V)$
	Ré-échantill.	0.0660	$p(A) \neq p(M)$	0.0212	$p(M) \neq p(V)$
	Régr. Log.	1.0000	$p(A) = p(M)$	0.0253	$p(M) \neq p(V)$

Tableau 15: Sommaire des *p-values* pour la Trace 6 selon les différentes techniques statistiques utilisées pour chaque comparaison simple.

I.2 Analyse univariée (Comparaisons multiples)

Les pages qui suivent présentent un résumé des résultats de la comparaison multiple $H_0 : p(A) = p(M) = p(V)$ versus $H_1 : p(A) \neq p(M)$ ou $p(M) \neq p(V)$ appliquée dans le cas d'une analyse univariée.

SOMMAIRE DES RESULTATS POUR LA TRACE 1			
Variable	Technique	A vs M vs V	Conclusion
B	Ré-échantill.	0.0002	H_1
	Régr. Log.	0.0000	H_1
P	Ré-échantill.	0.0016	H_1
	Régr. Log.	0.0000	H_1
I	Ré-échantill.	0.0040	H_1
	Régr. Log.	0.0000	H_1

Tableau 16: Sommaire des *p-values* pour la Trace 1 selon les différentes techniques statistiques utilisées.

SOMMAIRE DES RESULTATS POUR LA TRACE 2			
Variable	Technique	A vs M vs V	Conclusion
B	Ré-échantill.	0.0156	H_1
	Régr. Log.	0.0000	H_1
P	Ré-échantill.	0.0350	H_1
	Régr. Log.	0.0000	H_1
I	Ré-échantill.	0.9174	H_0
	Régr. Log.	0.0000	H_1

Tableau 17: Sommaire des *p-values* pour la Trace 2 selon les différentes techniques statistiques utilisées.

SOMMAIRE DES RESULTATS POUR LA TRACE 3			
Variable	Technique	A vs M vs V	Conclusion
<i>B</i>	Ré-échantill.	0.3834	H_0
	Régr. Log.	0.0000	H_1
<i>P</i>	Ré-échantill.	0.5698	H_0
	Régr. Log.	0.0642	H_1
<i>I</i>	Ré-échantill.	0.3940	H_0
	Régr. Log.	0.0000	H_1

Tableau 18: Sommaire des p -values pour la Trace 3 selon les différentes techniques statistiques utilisées.

SOMMAIRE DES RESULTATS POUR LA TRACE 4			
Variable	Technique	A vs M vs V	Conclusion
<i>B</i>	Ré-échantill.	0.0000	H_1
	Régr. Log.	0.0000	H_1
<i>P</i>	Ré-échantill.	0.0118	H_1
	Régr. Log.	0.1957	H_0
<i>I</i>	Ré-échantill.	0.0006	H_1
	Régr. Log.	0.0000	H_1

Tableau 19: Sommaire des p -values pour la Trace 4 selon les différentes techniques statistiques utilisées.

SOMMAIRE DES RESULTATS POUR LA TRACE 5			
Variable	Technique	A vs M vs V	Conclusion
<i>B</i>	Ré-échantill.	0.6840	H_0
	Régr. Log.	0.0000	H_1
<i>P</i>	Ré-échantill.	0.9208	H_0
	Régr. Log.	0.0205	H_1
<i>I</i>	Ré-échantill.	0.6954	H_0
	Régr. Log.	0.0205	H_1

Tableau 20: Sommaire des *p-values* pour la Trace 5 selon les différentes techniques statistiques utilisées.

SOMMAIRE DES RESULTATS POUR LA TRACE 6			
Variable	Technique	A vs M vs V	Conclusion
<i>B</i>	Ré-échantill.	0.0000	H_1
	Régr. Log.	0.0000	H_1
<i>P</i>	Ré-échantill.	0.0000	H_1
	Régr. Log.	0.0000	H_1
<i>I</i>	Ré-échantill.	0.0106	H_1
	Régr. Log.	0.0001	H_1

Tableau 21: Sommaire des *p-values* pour la Trace 6 selon les différentes techniques statistiques utilisées.

I.3 Résultats du test de McNemar

Le tableau qui suit présente un résumé des résultats (*p-values*) de l'analyse utilisant le test de McNemar.

SOMMAIRE DES RESULTATS DU TEST DE McNEMAR					
Trace #	Variable	A vs M	Conclusion	M vs V	Conclusion
Trace 1	B	0.00004	$p(A) \neq p(M)$	0.1573	$p(M) = p(V)$
	P	0.0047	$p(A) \neq p(M)$	0.0082	$p(M) \neq p(V)$
	I	0.7518	$p(A) = p(M)$	0.0001	$p(M) \neq p(V)$
Trace 2	B	0.00006	$p(A) \neq p(M)$	0.3173	$p(M) = p(V)$
	P	0.0320	$p(A) \neq p(M)$	0.0201	$p(M) \neq p(V)$
	I	0.3428	$p(A) = p(M)$	1.0000	$p(M) = p(V)$
Trace 3	B	1.0000	$p(A) = p(M)$	0.00006	$p(M) \neq p(V)$
	P	0.1797	$p(A) = p(M)$	0.0253	$p(M) \neq p(V)$
	I	0.6171	$p(A) = p(M)$	0.0114	$p(M) \neq p(V)$
Trace 4	B	0.00001	$p(A) \neq p(M)$	0.1004	$p(M) = p(V)$
	P	0.0008	$p(A) \neq p(M)$	0.3173	$p(M) = p(V)$
	I	0.0008	$p(A) \neq p(M)$	0.0201	$p(M) \neq p(V)$
Trace 5	B	0.0513	$p(A) = p(M)$	0.7518	$p(M) = p(V)$
	P	0.5839	$p(A) = p(M)$	0.5839	$p(M) = p(V)$
	I	0.5839	$p(A) = p(M)$	0.5839	$p(M) = p(V)$
Trace 6	B	0.00004	$p(A) \neq p(M)$	0.0082	$p(M) \neq p(V)$
	P	0.0023	$p(A) \neq p(M)$	0.0009	$p(M) \neq p(V)$
	I	0.0023	$p(A) \neq p(M)$	0.1797	$p(M) = p(V)$

Tableau 22: Sommaire des *p-values* pour chacun des tracés selon le test de McNemar. Ce test a été utilisé par un groupe de médecins de l'Hôpital Général d'Ottawa afin de vérifier si les techniques de contrôle de la condition foetale étaient équivalentes. Ces résultats seront comparés à ceux obtenus avec les différentes techniques statistiques employées dans le présent document.

Bibliographie

- [1] BICKEL, Peter J., DOKSUM, Kjell A., *Mathematical Statistics: Basic Ideas and Selected Topics*, Californie, Holden-Day Inc., 1977.
- [2] BISHOP, Yvonne M., FIENBERG, Stephen E., HOLLAND, Paul W., *Discrete Multivariate Analysis: Theory and Practice*, The MIT Press, 1991.
- [3] GOOD, Phillip, *Permutation Tests: A practical Guide to Resampling Methods for Testing Hypotheses*, Springer-Verlag, 1993.
- [4] HOSMER, David W., LEMESHOW, Stanley, *Applied Logistic Regression*, New York, John Wiley & Sons Inc, 1989.
- [5] JOHNSON, Richard A., WICHERN, Dean W., *Applied Multivariate Statistical Analysis*, 3^e édition, New Jersey, Prentice-Hall Inc., 1992.
- [6] SIMPSON, Nigel, OPPENHEIMER, Lawrence W., SIREN, Andrew, McDONALD, Omid, McDONALD, David, DABROWSKI, André, **Accuracy of Strategies for Monitoring Fetal Heart Rate in Labor**, 1995, "preprint".
- [7] WESTFALL, Peter H., YOUNG, S. Stanley, *Resampling-Based Multiple Testing: Examples and Methods p-value adjustment*, New York, John Wiley & Sons Inc, 1993.