

Comparaison du niveau de réponse immunitaire moyen entre des groupes de traitements comprenant des répondeurs et des non répondeurs

Projet Sanofi Pasteur SEME 2017 Lyon

Aya EL DAKDOUKI * Caroline ROBOT †
Bertrand SIMON ‡ Cécile TAING §

3 Février 2017

1 Introduction

Dans le cadre du développement de nouveaux vaccins, les étapes pré-cliniques consistent en la comparaison des données expérimentales de réponses immunitaires obtenues après injection de candidats vaccins. Ces candidats vaccins diffèrent en général selon le type d'antigène, d'adjuvant et la formulation du vaccin. L'objectif est de déterminer par la comparaison des réponses quantifiées le candidat vaccin qui générerait la meilleure réponse immunitaire. En général, les valeurs des réponses immunitaires déterminées sont *censurées à gauche* en raison de la limite de quantification de l'appareil de mesure utilisé. En dessous de ce *seuil* de quantification spécifique à un appareil de mesure (Limit of quantification ou LOQ), les données sont considérées comme très peu fiables: on peut considérer que la seule information disponible est le fait que ces données sont plus petites que le seuil LOQ. Ces individus sont dénommés *non répondeurs*, ce qui signifie que leurs réponses ne sont pas mesurées, sans être forcément nulles. Afin de les traiter, leurs réponses sont pour l'instant remplacées par la valeur du seuil LOQ divisée par deux. Ainsi, l'objet de ce travail proposé par Sanofi-Pasteur est de comprendre comment utiliser toute l'information disponible, à

*Université de Lille 1, Laboratoire Paul-Painlevé UMR CNRS 8524, Bâtiment M2, Cité Scientifique, 59655 Villeneuve-d'Ascq, Lille, France et Université de Lorraine, Laboratoire de la Fédération Charles Hermites (FR 3198 du CNRS) Equipe ABC, Nancy, France. Email: aya.dakdouki@live.com

†Université de Nantes, Laboratoire de Mathématiques Jean Leray, 2 rue de la Houssinière, BP 92208, F-44322 Nantes Cedex 3, France. Email : caroline.robet@univ-nantes.fr

‡LIP, ENS Lyon, 46 allée d'Italie, Lyon, France. Email: bertrand.simon@ens-lyon.fr

§Sorbonne Universités, UPMC Univ Paris 06, Inria, Laboratoire Jacques-Louis Lions UMR CNRS 7598, F75005 Paris, France. Email: cecile.taing@upmc.fr

savoir les réponses immunitaires des répondeurs et non-répondeurs, pour construire un modèle statistique permettant la comparaison des différents groupes de traitement.

Nous proposons ici des modèles statistiques de censure et de mélange de lois pour modéliser les réponses immunitaires des sujets. Trois problématiques ont été identifiées en fonction des jeux de données fournis. Nous exposons dans la suite les trois approches utilisées pour traiter ces problématiques.

2 Présentation du problème

L'objectif commun à chaque problématique est de proposer un protocole de test qui permette de comparer deux vaccins ou plus, en fonction des réponses obtenues par chacun sur un groupe de sujets. La difficulté principale, comme exposé plus haut, est que l'appareil de mesure utilisé n'est pas précis sur des réponses faibles (en dessous d'un seuil), pour lesquelles on ne connaît donc pas la valeur exacte, mais seulement la présence dans un intervalle $[0; \text{seuil}]$. Ce type de données est connu dans la littérature comme une distribution *censurée à gauche*. Un objectif préliminaire sera de modéliser la distribution statistique des réponses obtenues, avant de déterminer par exemple si les différences obtenues dans l'estimation des paramètres de deux vaccins sont significatives.

Dans toute cette étude, on suppose pour simplifier qu'il n'existe pas de sujets ayant une réponse totalement nulle. En pratique, certaines méthodes peuvent permettre de détecter sous certaines conditions de tels sujets.

Une deuxième hypothèse que nous considérons est que dans une population *homogène*, la distribution de la réponse immunitaire pour un vaccin donné suit une loi normale¹, dont la variance ne dépend pas du vaccin. Un exemple typique de population homogène est un groupe de souris utilisé en étude pré-clinique, qui sont donc censées avoir des systèmes immunitaires *équivalents*.

3 Étude pré-clinique sur des groupes de souris

Dans ce premier problème, nous nous intéressons à une étude pré-clinique de plusieurs vaccins sur des groupes de souris. L'une des difficultés inhérentes à ce type d'étude étant le faible effectif des sujets dans chaque groupe de traitement.

Nous choisissons de supposer que la réponse immunitaire suit une loi gaussienne, comme précisé plus haut. Nous avons considéré l'étude dont les résultats sont résumés en Figure 1. Les mesures obtenues en dessous du seuil de quantification sont indiquées par "--". L'objet de cette section est donc de trouver les estimations paramétriques de la distribution gaussienne censurée qui correspondrait aux données obtenues, et de déterminer s'il existe des différences significatives entre les groupes pour comparer l'efficacité des candidats vaccins.

¹nous définissons la réponse comme étant le logarithme de la concentration mesurée, d'où l'hypothèse d'une loi normale et non log-normale

A	B	C	D	E	F	G	H	I	J	K
--	1.32	--	1.7	1.26	--	--	--	--	--	--
--	--	1.74	--	--	1.41	1.69	--	1.36	--	--
1.92	1.23	--	--	1.34	--	--	--	--	--	2.54
--	--	1.3	--	--	--	--	1.32	--	1.68	--
--	--	1.96	--	--	--	--	1.38	--	--	1.41
2.11	2.08	2.17	--	--	--	1.72	--	--	--	--
--	2.25	1.23	1.92	1.2	--	--	1.91	1.81	1.71	--
1.41	1.28	1.38	1.23	--	--	--	--	--	1.34	--

Figure 1: Un exemple de données utilisées: 11 groupes de 8 souris avec un seuil égal à 1.15.

3.1 Estimation de la moyenne de chaque échantillon

Le premier objectif est d'estimer la moyenne de la loi normale propre à chaque échantillon.

Précédemment, et il s'agit précisément de la problématique qui a motivé ces travaux, les non-répondeurs n'étaient pas pris en compte de manière satisfaisante. Une méthode qui pouvait être suivie renvoie simplement la moyenne obtenue sur les répondeurs, en écartant totalement les non-répondeurs. L'hypothèse sous-jacente étant que la proportion de non-répondeurs n'est pas significative, ce qui n'est pas vérifié en pratique. Une deuxième méthode consiste à remplacer la valeur de chaque non-répondeur par la moitié du seuil, puis de calculer la moyenne sur tous les échantillons. Le problème ici est que la moyenne des non-répondeurs dépend de la distribution de la loi, et n'est donc pas égale à la moitié du seuil.

Afin de prendre en compte l'information donnée par les non-répondeurs dans sa totalité, la méthode proposée détermine les paramètres de la loi normale maximisant la vraisemblance sur un groupe donné. Une implémentation de cet estimateur prenant en compte les non-répondeurs, c'est-à-dire une censure à gauche, est donnée par la fonction `fitdistcens` du package R `fitdistrplus`. On peut noter que cette fonction est plus générale, car elle permet également de traiter des censures par intervalles, et non seulement à gauche.

Avant d'utiliser cette estimation pour comparer plusieurs groupes de sujets, nous avons comparé la précision de cette méthode par rapport aux précédentes, sur des données simulées.

Nous avons simulé aléatoirement 20 groupes de 100 individus selon une même loi normale arbitraire, de moyenne 15 et d'écart-type 5. Afin de simuler la censure à gauche introduite par un appareil de mesure, nous avons introduit un seuil, qui prendra dans le protocole plusieurs valeurs comprises entre 5 et 16. A seuil fixé, la valeur obtenue pour les individus à gauche du seuil est *masquée*, ou censurée à gauche: on retient seulement que cette valeur est plus petite que le seuil. Toujours pour un seuil donné, on estime la moyenne de chaque groupe via chaque méthode, et on calcule l'erreur moyenne commise par chaque méthode. Il reste enfin à tracer l'erreur commise par chaque méthode en fonction de la

valeur du seuil pour obtenir les résultats présentées sur la Figure 2.

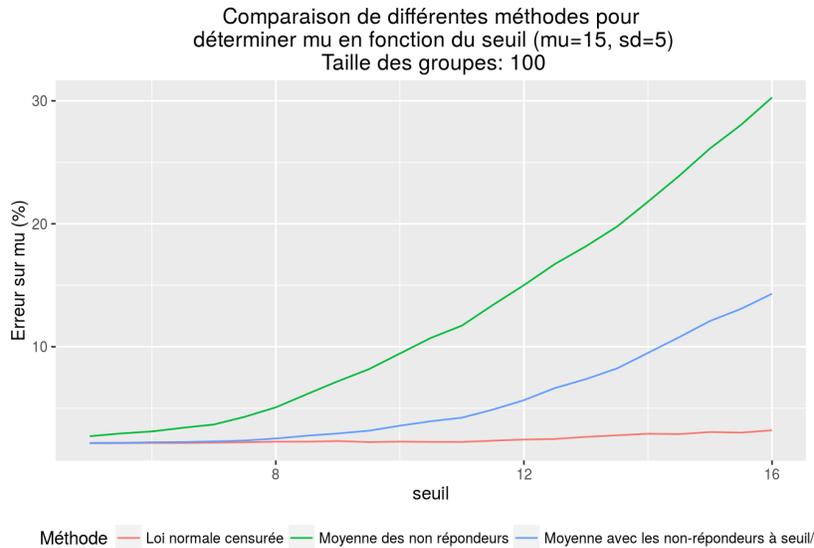


Figure 2: Comparaison des erreurs commises dans l’estimation de la moyenne par différentes méthodes sur des groupes simulés de 100 individus.

Sur cette figure, la méthode proposée est tracée en rouge. On peut tout d’abord remarquer que les trois méthodes commettent peu d’erreur lorsque le seuil est petit devant la moyenne, ce qui est attendu car le nombre de non-répondeurs est très faible. L’enseignement principal de ces courbes est que la méthode proposée ne dépasse pas quelques pourcents d’erreurs sur l’estimation de la moyenne, même lorsque le seuil est au-dessus de la moyenne réelle. Elle est donc beaucoup plus robuste que les méthodes précédentes, qui atteignent respectivement 14% et 30% d’erreur pour un tel seuil.

La Figure 3 représente les mêmes résultats pour des groupes de taille 10. Les conclusions sont similaires, avec comme l’on peut s’y attendre, une précision plus faible pour chaque méthode, sans toutefois dépasser les 10% d’erreur pour la méthode proposée.

Nous proposons donc un moyen robuste d’estimer la réponse moyenne d’un vaccin grâce aux mesures effectuées sur un petit nombre d’individus.

3.2 Comparaison entre différents groupes

L’objectif principal est maintenant de déterminer si les écarts entre les moyennes estimées de plusieurs groupes sont statistiquement significatifs. Plus précisément, on cherche à définir un protocole de test qui prend en entrée les réponses obtenues sur plusieurs groupes de sujets et qui détermine si l’on peut raisonnablement affirmer que ces groupes ont des moyennes différentes. Ce résultat est

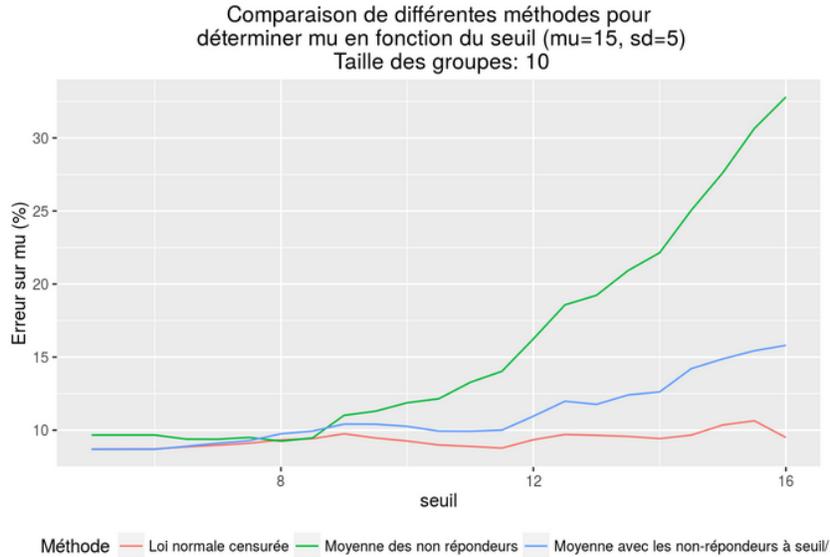


Figure 3: Comparaison des erreurs commises dans l'estimation de la moyenne par différentes méthodes sur des groupes simulés de 100 individus.

donné sous la forme d'une p -valeur (ou p -value en anglais). Si cette p -valeur est trop éloignée de 0, le test ne permettra pas de conclure qu'il existe des différences significatives.

Le premier test que nous avons utilisé permet uniquement de traiter deux groupes de sujets. Il s'agit d'une variation du classique test de Student, qui utilise les paramètres estimés (moyenne et écart-type) ainsi que le nombre de sujets dans chaque groupe. Cette variante nous est utile car nous ne connaissons pas précisément la réponse des non-répondeurs. Ce test est implémenté dans la fonction `tsum.test` du package R `PASWR`.

Le deuxième test que nous avons utilisé permet de considérer plusieurs groupes de sujets, et de déterminer si l'on peut affirmer que tous les groupes ne suivent pas une loi normale de même moyenne. Ce test consiste en une variante de l'analyse de variance (ANOVA) classique, prenant en compte le fait que certaines données ne sont pas connues précisément, mais sont censurées à gauche (on sait qu'elles sont inférieures à une valeur seuil). Ce test est implémenté dans la fonction `cenmle` du package R `NADA`.

Nous avons utilisé ces tests sur les données présentées en Figure 1. Les seules paires de groupes qui sont significativement différents avec une p -valeur inférieure à 0.05 sont les paires C-F, B-F, E-F, C-G et C-I. Si l'on considère plusieurs groupes, seuls les groupes de faible taille contenant une de ces paires donnent également une p -valeur inférieure à 0.05. Une conclusion que l'on peut tirer est donc que, si l'on avait mesuré seulement deux groupes de sujets ayant respectivement les réponses des groupes C et F, on conclurait à des moyennes

significativement différentes. Seulement, étant donné que ces groupes ont été choisis parmi une étude de onze groupes, la p-valeur obtenue par cette comparaison ne peut être interprétée comme telle. En effet, le test effectué sur l'ensemble des groupes renvoie une forte p-valeur, et l'on ne peut donc pas conclure que les groupes de sujets ne suivent pas tous la même loi. La différence observée pourrait être due uniquement à l'aléatoire.

4 Étude clinique sur deux groupes humains

Nous étudions les données des réponses immunitaires de deux groupes de sujets pour deux vaccins différents, nommés vaccin 1 et vaccin 2.

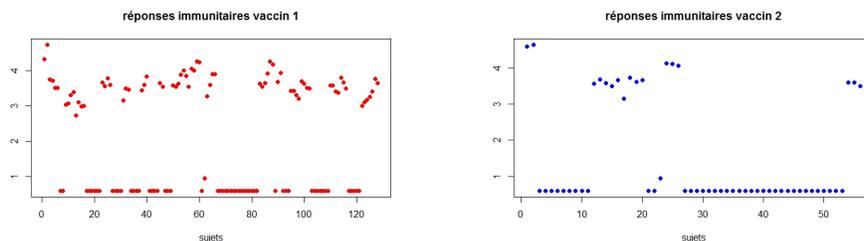


Figure 4: Données de 2 vaccins

L'analyse de variance montre bien que la distribution des données n'est pas gaussienne. L'idée est donc de modéliser la distribution de chacun des vaccins par une combinaison convexe de lois gaussiennes, de la forme :

$$X = pZ_1 + (1 - p)Z_2, \quad 0 < p < 1,$$

où Z_1 et Z_2 suivent des lois gaussiennes indépendantes $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$. Les deux groupes étant bien séparés dans le problème présent, il est raisonnable d'estimer les poids par la fréquence d'individus dans les groupes fortement répondeurs et faiblement/non répondeurs. Cependant, les données comportent une grande proportion de valeurs censurées, il n'est donc pas possible d'estimer les paramètres de Z_1 . Nous choisissons donc de fixer la loi Z_1 , de générer des observations de Z_1 de telle sorte qu'une partie des valeurs soient au-dessus du seuil et de ramener au seuil toutes les valeurs inférieures au seuil. Le but est de pouvoir tester notre modèle dans le cas où l'on a quelques valeurs faibles mais au-dessus du seuil, contrairement à cet exemple particulier où un seul sujet est faiblement répondeur, c'est-à-dire dont la réponse immunitaire est proche du seuil de quantification. L'approximation des données non censurées appartenant à la distribution Z_2 est effectuée avec la fonction `fidist`, celle des données générées et censurées de la loi Z_1 avec `fitdistcens`. Les résultats obtenus sont les suivants pour les estimations de $(p, \mu_1, \sigma_1 ; 1 - p, \mu_2, \sigma_2)$:

- Estimations vaccin 1 : (0.45, 0.7, 0.37 ; 0.55, 3.59, 0.35)

- Estimations vaccin 2 : (0.69, 0.64, 0.38 ; 0.31, 3.79, 0.38)

On souhaite ensuite comparer ces vaccins pour déterminer s'ils sont significativement différents ou si les différences ne sont dues qu'à l'aléa des observations. Un premier test permet de comparer la proportion p de non/faiblement répondeurs entre les deux vaccins. Ce test est basé la fonction `prop.test` du package `stats`. Dans notre exemple, on trouve que la proportion de non-répondeurs du vaccin 2 est significativement plus grande que la proportion du vaccin 1 (p-valeur de 0,003 dans le test d'égalité). Un autre test possible dans ce modèle est le test de Student qui permet de comparer la réponse moyenne μ_2 des fortement répondeurs pour chacun des vaccins (`t.test` ou `tsum.test` avec le package `PASWR`). Dans notre exemple, l'hypothèse d'égalité des moyennes pour les gaussiennes hautes n'est pas rejetée (la p-valeur de ce test est à 0,067). Suivant la maladie concernée, on pourra ensuite privilégier un vaccin qui donne une réponse immunitaire forte (quitte à se contenter d'un taux plus faible de répondeurs) ou alors préférer un vaccin avec un très faible taux de non répondeurs (quitte à ce que la moyenne ne soit pas forcément très élevée).

Pour une comparaison de multiples vaccins, on peut faire des tests similaires sur p (fonction `prop.test` avec plusieurs groupes) et sur μ_2 (ANOVA puis tests post-hoc).

Dans ce cas d'un mélange de lois avec deux groupes bien distincts, on a ainsi mis en avant des techniques qui nous permettent d'estimer les paramètres des lois et qui permettent de comparer plusieurs vaccins avec ce type de réponse immunitaire.

5 Étude sur le cas d'une grippe non saisonnière

La Figure 5 montre la répartition de la réponse immunitaires de 192 sujets avant et après vaccination contre la grippe. Les données avant vaccination n'ont pas

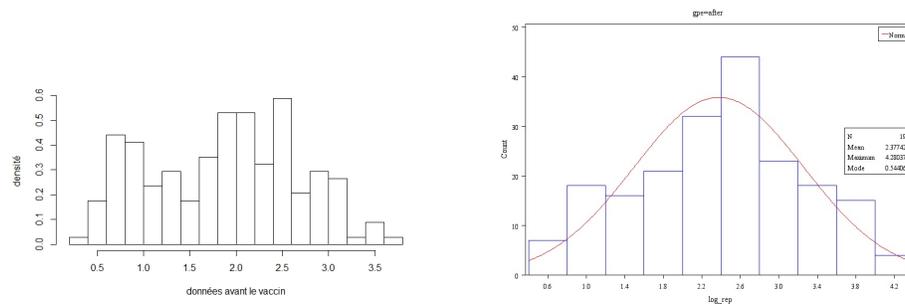


Figure 5: Données avant et après vaccination contre la grippe

été étudiées car les sujets n'ont pas tous été confrontés le même nombre de fois à la grippe et ont ainsi une réponse immunitaire non homogène. Après le vaccin, on propose comme modèle de répartition le mélange de deux lois gaussiennes

censurées à gauche Z_1 et Z_2 .

$$X = pZ_1 + (1 - p)Z_2,$$

On estime les paramètres $(p, \mu_1, \sigma_1, \mu_2, \sigma_2)$ de la combinaison convexe grâce à la fonction `mixcensoredInt` (mise en ligne par S. Gelissen). Cette fonction utilise l'algorithme itératif EM (Expectation Maximization). Cet algorithme permet d'estimer les paramètres par maximum de vraisemblance dans les cas où la vraisemblance est analytiquement impossible à maximiser, ou bien lorsque l'on a des données manquantes. Il est donc particulièrement adapté aux modèles de mélange de lois gaussiennes. Sur le cas des 192 sujets de l'exemple, on obtient les estimations suivantes: répartition $p = 1.5\%$ et $(1 - p) = 98.5\%$, paramètres gaussienne à gauche: $\mu_1 = 1.53$, $\sigma_1 = 0.79$, paramètres gaussienne à droite: $\mu_2 = 2.76$, $\sigma_2 = 0.66$.

La proportion p estimée de personnes dans la gaussienne à gauche censurée est faible. On peut se demander si un modèle avec une seule gaussienne serait plus en adéquation avec les données. Notre modèle de mélange n'a pas pu être testé. Cependant, si on se base sur l'article [1], on pourrait vérifier notre hypothèse en réalisant un test de rapport de vraisemblance. Sous l'hypothèse nulle, on suppose que la loi de l'échantillon est une loi normale censurée (i.e $p = 0$ ou $p = 1$) que l'on teste contre l'hypothèse alternative où la loi de l'échantillon est un mélange de deux lois normales censurées (i.e $0 < p < 1$). On obtient une région de rejet de l'hypothèse nulle à 95% sous la forme $\{-2 \log(R) > 2, 71\}$, où

$$R = \frac{\max_{(p, \mu, \sigma) \in \{1\} \times \mathbb{R}^+ \times \mathbb{R}} \mathcal{V}(x_1, \dots, x_n | p, \mu, \sigma)}{\max_{(p, \mu, \tilde{\mu}, \sigma, \tilde{\sigma}) \in [0, 1] \times \mathbb{R}^+ \times 2 \times \mathbb{R}^2} \mathcal{V}(x_1, \dots, x_n | p, \mu, \tilde{\mu}, \sigma, \tilde{\sigma})}.$$

Dans notre exemple, le numérateur est bien estimé car il s'agit de la vraisemblance d'une loi gaussienne censurée à gauche mais le dénominateur nécessite d'optimiser la fonction de vraisemblance sur de nombreux paramètres ce qui n'a pas pu être fait de manière suffisamment précise.

Dans le cas de la grippe, une idée de poursuite d'étude consisterait à prendre en compte la réponse immunitaire de chaque patient avant vaccination. On pourrait regarder des couples d'observations qui correspondraient à la réponse immunitaire avant/après pour chaque sujet pour estimer l'efficacité du vaccin.

6 Conclusion et perspectives

Nous avons utilisé différentes hypothèses pour analyser les données de réponse immunitaire selon leur taille et leur répartition et proposé des méthodes d'estimation des paramètres et de calcul d'erreur. L'hypothèse de normalité pour la distribution de la réponse immunitaire peut paraître forte, en particulier dans la première étude des les groupes de souris. Une question intéressante pourrait être de déterminer à partir de quel effectif minimal l'hypothèse de normalité deviendrait cohérente et permettrait une comparaison performante des différents candidats vaccins.

Par ailleurs, un modèle est proposé dans l'article [1] et utilise des données provenant d'une étude sur le vaccin contre la rougeole. Les observations sont modélisées par un couplage d'une loi normale, d'une distribution sous le seuil et d'une loi de Bernoulli. L'approche proposée dans cet article diffère de la nôtre dans le choix de la description des réponses négatives des sujets. Les réponses négatives peuvent être issues soit de la loi normale censurée, soit de la distribution en dessous du seuil de quantification. La provenance des réponses négatives est modélisée par une loi de Bernoulli. La distribution en dessous du seuil est inconnue. On peut de nouveau appliquer un test de rapport de vraisemblance pour déterminer si le modèle possède une distribution sous le seuil. Un algorithme utilisant le maximum de vraisemblance est développé pour l'estimation des paramètres des lois impliquées dans ce modèle Bernoulli/normale. Une question à étudier pourrait être la comparaison de cette méthode à celle que nous avons développée dans ce présent rapport.

Remerciements

Nous remercions l'AMIES pour le financement de la SEME de Lyon 2017. Nous remercions également Aymeric de Montfort, biostatisticien de Sanofi-Pasteur, qui nous a proposé ce sujet et qui a pris le temps pour répondre à toutes nos questions au cours de cette semaine, ainsi que les organisateurs de de la SEME, Fabien Crauste, Céline Helbert, Véronique Maume-Deschamps, Céline Vial et François Wahl, pour toute l'aide apportée dans notre travail.

References

- [1] MOULTON, L. H. AND HALSEY, N. A., *A mixture model with detection limits for regression analyses of antibody response to vaccine*. BIOMETRICS 51, 1995, 1570-1578.