

---

**Translations of Abstracts**

---

**Discussion Paper**

H. Chen, Z. Geng, and X.-H. Zhou

675

*Identifiability and Estimation of Causal Effects in Randomized Trials with Noncompliance and Completely Nonignorable Missing Data*

Dans ce papier nous considérons d'abord l'identifiabilité des paramètres dans des essais cliniques randomisés avec noncompliance et réponses manquantes. Nous montrons que sous certaines conditions les paramètres d'intérêt sont identifiables même dans différents types de données manquantes non-ignorables : c'est-à-dire, le mécanisme induisant une donnée manquante dépend de la réponse. Nous dérivons alors leurs estimateurs du maximum de vraisemblance et de la méthode des moments et évaluons par des simulations leurs propriétés sur des échantillons finis en terme de biais, d'efficacité et de robustesse. Notre analyse de sensibilité montre que le modèle choisi pour les données manquantes non-ignorables influe fortement sur le paramètre d'effet causal moyen de compliance (CACE). Notre nouvelle méthode fournit de nouveaux et utiles modèles pour données manquantes non-ignorables par rapport au modèle existant d'ignorable latent, qui garantissent l'identifiabilité des paramètres pour estimer le CACE dans un essai clinique randomisé avec noncompliance et données manquantes.

**Biometric Methodology**

D.B. Boos, L.A. Stefanski, and Y. Wu

692

*Fast FSR Variable Selection with Applications to Clinical Trials*

Nous développons une nouvelle version de la méthode FSR (False Selection Rate) de sélection de variables, de Wu, Boos et Stefanski (2007) ne nécessitant pas de simulation. Cette version permet une estimation du paramètre de réglage de la sélection ascendante par simple calcul manuel à partir d'une table de résumés, même dans les situations où le nombre de variables explicatives est plus élevé que la taille de l'échantillon. En raison de la simplicité des calculs, la méthode peut être utilisée dans des tests de permutation et dans des boucles d'agrégation (bagging loops) pour améliorer les prédictions. Nous illustrons par des essais cliniques avec la régression linéaire, la régression logistique, et le modèle de régression de Cox.

Nous nous intéressons à la sélection de variables en classification non supervisée. Nous envisageons le problème comme un problème de sélection de modèles de mélanges de lois de probabilité. Pour cela, nous généralisons le modèle de Raftery et Dean (2006) en spécifiant mieux le rôle de chaque variable. Notre modèle ne nécessite aucune hypothèse a priori sur le lien entre les variables sélectionnées et celles écartées pour la classification. Les modèles sont comparés grâce au critère BIC. Le statut des variables est obtenu grâce à un algorithme imbriquant deux algorithmes de sélection descendante avec remise en cause de variables pour la classification et la régression linéaire. Nous établissons l'identifiabilité des modèles et démontrons la consistance du critère de sélection sous des conditions de régularité. Nous mettons en évidence l'intérêt de cette procédure de sélection de variables à l'aide d'exemples numériques sur des données simulées et sur une application génomique.

X. Huang and J.M. Tebbs 710  
*On Latent-variable Model Misspecification in Structural Measurement Error Models for Binary Response*

Nous considérons des modèles structuraux à erreur de mesure pour réponse binaire. Nous montrons que les estimateurs basés sur la vraisemblance obtenus en ajustant des modèles avec des réponses binaires poolées peuvent être beaucoup moins robustes aux erreurs de mesures sur les covariables en présence d'une mauvaise spécification du modèle de variable latente que les estimateurs correspondants pour réponses individuelles. De plus, en dépit de la perte d'information, pooler peut fournir des estimateurs des paramètres améliorés en termes d'erreur de mesure quadratique. Basés entre autre sur ces considérations, nous créons une nouvelle méthode diagnostique pour détecter la mauvaise spécification du modèle de variable latente dans les modèles structuraux à erreur de mesure avec une réponse binaire individuelle. Nous utilisons des simulations et des données de l'étude cardiologique de Framingham pour illustrer nos méthodes.

X. Huang, L.A. Stefanski, and M. Davidian 719  
*Latent-model Robustness in Joint Models for a Primary Endpoint and a Longitudinal Process*

La modélisation conjointe d'une réponse principale et d'un processus longitudinal via des effets aléatoires communs est largement utilisée dans beaucoup de domaines d'application. L'inférence basée sur la vraisemblance de modèles conjoints nécessite de spécifier les effets aléatoires du modèle. Une spécification inappropriée de ceux-ci peut remettre en question l'inférence. Nous présentons des méthodes pour diagnostiquer une mauvaise spécification des effets aléatoires qui conduit à une inférence incorrecte des modèles conjoints. Les méthodes sont illustrées à partir de données simulées et à l'aide de l'utilisation de données provenant d'une étude sur la densité osseuse chez des femmes en péri-ménopause et d'un essai clinique sur HIV.

*Estimation in Semiparametric Transition Measurement Error Models for Longitudinal Data*

Nous considérons des modèles de transition semi-paramétriques avec erreurs de mesures pour des données longitudinales, où une des covariables est mesurée avec erreur dans des modèles de transition, et aucune hypothèse de distribution n'est faite pour la distribution sous-jacente de la covariable non observée. Une approche d'équations d'estimation basée sur une méthode de pseudo-score conditionnel est proposée. Nous montrons que les estimateurs des coefficients de régression sont consistants et asymptotiquement normaux. Nous discutons du problème de la perte d'efficacité. Des études de simulations sont conduites pour examiner les performances de nos estimateurs dans un cadre fini. Les données longitudinales de l'étude des coûts et des services liés au HIV sont analysées comme illustration.

*A Semiparametric Joint Model for Longitudinal and Survival Data with Application to Hemodialysis Study*

Dans de nombreuses études cliniques longitudinales, le niveau et le taux de progression de biomarqueurs mesurés de façon répétitive sur chaque sujet permettent de quantifier la sévérité de la maladie et la susceptibilité du sujet à sa progression. Il est intéressant, sur les plans clinique et scientifique, d'associer ces mesures à un temps écoulé terminal clinique plus tardif tel que la survie du patient. Ceci est fait usuellement avec un modèle à paramètre partagé. Dans ces modèles, les données longitudinales de biomarqueurs et la survie de chaque sujet sont supposées conditionnellement indépendantes, étant donné son niveau de sévérité ou de susceptibilité (également appelé fragilité en termes statistiques). Dans cet article, nous étudions le cas où la distribution conditionnelle des données longitudinales est modélisée par un modèle linéaire à effets mixtes, et la distribution conditionnelle des données de survie est donnée par un modèle des hasards proportionnels de Cox. Nous admettons des coefficients de régression inconnus et des covariables dépendantes du temps dans les deux modèles. Les estimateurs proposés maximisent une correction exacte de la log-vraisemblance jointe avec les paramètres de fragilité éliminés comme paramètres de nuisance, une idée provenant de la correction de l'erreur de mesure des covariables dans les modèles à erreurs de mesure. On montre que la log-vraisemblance jointe corrigée est asymptotiquement concave et conduit à des estimateurs consistants et asymptotiquement gaussiens. Contrairement à la plupart des méthodes publiées pour une modélisation jointe, la méthode d'estimation proposée ne repose pas sur des hypothèses de distribution pour les paramètres de fragilité. La méthode proposée est étudiée dans des simulations et appliquée à un ensemble de données de l'étude Hémodialyse (HEMO).

*Semiparametric Transformation Models with Random Effects for Joint Analysis of Recurrent and Terminal Events*

Nous proposons une large classe de modèles de transformations semi-paramétriques avec des effets aléatoires pour l'analyse conjointe d'événements récurrents et d'un événement terminal. Les modèles de transformation incluent des modèles à risques / intensités proportionnels ainsi que des modèles à côtes (« odds ») proportionnels. On estime les paramètres du modèle par la méthode non-paramétrique du maximum de vraisemblance. On montre que les estimateurs sont cohérents, asymptotiquement normal et asymptotiquement efficace. Des algorithmes numériques simples et stables sont fournis pour calculer les estimations des paramètres et de leurs variances. Une importante étude de simulations démontre que les procédures proposées se comportent bien dans des situations réalistes. Des applications à deux études sur le VIH / AIDS sont présentées.

Q. Pan and D.E. Schaubel

753

*Flexible Estimation of Differences in Treatment-specific Recurrent Event Means in the Presence of a Terminating Event.*

Dans cet article, nous considérons le cas où l'évènement d'intérêt peut se produire de manière répétée chez le même sujet (i.e., un évènement récurrent; e.g., une hospitalisation) et peut être interrompu définitivement par un évènement terminal (e.g., le décès). Parmi les différentes manières pour modéliser les données avec des évènements récurrents / terminaux, la moyenne marginale (i.e., moyennant sur la distribution de survie) est d'un intérêt majeur dans les perspectives en santé publique et en économie de la santé. Le plus souvent, la différence entre les moyennes des évènements récurrents spécifiques au traitement ne sera pas constante dans le temps, en particulier, quand des différences spécifiques aux traitements en terme de survie existent. Dans une telle situation, il paraît plus raisonnable de quantifier l'effet du traitement fondé sur la différence cumulée de la moyenne d'un évènement récurrent, par opposition avec la différence instantanée du taux. Nous proposons une méthode qui compare les traitements par des estimations séparées des probabilités de survie et des taux d'évènements récurrents conditionnel à la survie. Le nombre moyen d'évènements est ensuite obtenu en intégrant les estimations. La méthode proposée combine un modèle additif pour le taux d'évènements récurrents conditionnel et un modèle à risques proportionnels pour le risque associé à l'évènement terminal. Les effets du traitement sur la survie et sur le taux d'évènements récurrents parmi les survivants sont estimés dans la construction de notre mesure et permet une explication du mécanisme qui génère la différence à l'étude. L'exemple qui a motivé cette recherche est l'occurrence répétée d'hospitalisations parmi des bénéficiaires de greffes de rein, où notre intérêt porte sur l'effet de Donneurs à Critères Etendus (ECD) comparé aux greffes de rein de Donneurs à Critères non-Etendus (non-ECD) sur le nombre moyen d'hospitalisation.

M. De Iorio, W.O. Johnson, P. Mueller and G.L. Rosner

762

*Bayesian Nonparametric Non-Proportional Hazards Survival Modelling*

Nous développons un modèle de Processus de Dirichlet Dépendant pour l'analyse de données de survie. Une caractéristique majeure de l'approche proposée est qu'il n'est pas nécessaire que les courbes de survie estimées obtenues satisfassent l'hypothèse des hasards proportionnels. Nous donnons une illustration basée sur un essai clinique de

cancérologie, où les probabilités de survie estimées aux temps précoces de l'étude sont inférieures pour les sujets traités par doses élevées par rapport aux sujets traités par faibles doses, alors que la situation est inversée aux temps plus tardifs de l'étude, ceci pouvant être dû à l'effet toxique des doses élevées pour les sujets en mauvaise condition au début de l'étude.

B. Scarpa and D.B. Dunson

772

*Bayesian Hierarchical Functional Data Analysis Via Contaminated Informative Priors*

Des approches flexibles variées ont été proposées pour l'analyse de données fonctionnelles, permettant à la fois une courbe moyenne et une distribution autour de la moyenne inconnues. De telles méthodes sont les plus utiles lorsqu'on dispose d'une information a priori limitée. Motivé par des applications à la modélisation des courbes de température pendant le cycle menstruel cet article propose une approche flexible pour incorporer une information a priori dans des analyses bayésiennes semi-paramétriques de données fonctionnelles hiérarchiques. L'approche proposée se base sur la spécification de la distribution des fonctions comme un mélange d'un modèle hiérarchique paramétrique et d'une contamination non-paramétrique. La composante paramétrique est choisie en se basant sur la connaissance a priori, tandis que la contamination est caractérisée par un processus fonctionnel de Dirichlet. Dans l'application motivante, la composante de contamination permet des formes de courbes non prévues pour des cycles menstruels en contexte pathologique. Les méthodes sont développées pour les calculs a posteriori, et nous appliquons notre approche à des données d'une étude européenne sur la fécondabilité.

D. Senturk, D. Nguyen, F. Tassone, R.J. Hagerman, R.J. Carroll and P.J. Hagerman 781

*Covariate Adjusted Correlation Analysis with Application to FMR1 Premutation Female Carrier Data*

Afin de pouvoir analyser des données moléculaires obtenues chez des femmes porteuses de prémutation au sein du gène de l'X fragile (FMR1, Fragile X mental retardation), nous proposons une nouvelle méthode d'analyse de corrélation permettant d'ajuster sur des covariables afin d'étudier l'association entre les taux d'ARN messager (ARNm) et le nombre d'expansion de triplets CGG dans le gène FMR1. L'étude de l'association entre les variables moléculaires chez les femmes porteuses nécessite d'ajuster sur le ratio d'activation (ActRatio) qui est une mesure qui prend en compte les effets protecteurs du chromosome X normal de ces femmes. Cependant, il existe des incertitudes quant à l'effet exact de l'ActRatio sur les mesures moléculaires étudiées. Afin d'en tenir compte, nous développons une méthode flexible d'ajustement qui traite de manière non paramétrique les effets additifs et multiplicatifs de l'ActRatio. La méthode proposée utilise des corrélations conditionnelles locales pour estimer, à partir de la méthode locale des moments, le coefficient de corrélation de Pearson entre deux variables tout en ajustant pour une troisième variable observée. Les estimateurs obtenus par la méthode locale des moments sont moyennés pour obtenir un estimateur final de la corrélation ajustée, estimateur dont la consistance est démontrée. Nous proposons aussi un test d'adéquation de la modélisation non paramétrique des effets additifs et multiplicatifs. L'efficacité de la

méthode est illustrée par des études de simulations. Son application aux données sur les prémutations de FMR1 chez 165 femmes porteuses montre que l'association entre les taux d'ARNm et les triplets CGG est renforcée après ajustement sur l'ActRatio. Finalement, ces résultats fournissent des arguments supplémentaires à ceux déjà décrits dans la littérature en faveur de l'ajustement simultané sur les effets additifs et multiplicatifs de l'ActRatio.

D. Telesca, L.Y.T. Inoue, M. Neira, R. Etzioni, M. Gleave and C. Nelson 793  
*Differential Expression and Network Inferences through Functional Data Modeling*

Les données de puces longitudinales consistent en des expressions d'ARN messenger (ARNm) d'un jeu de gènes mesurés à différents points temporels. De telles données sont supposées refléter les processus biologiques sous-jacents qui se développent dans le temps. Dans cet article, nous proposons un modèle qui nous permet d'examiner l'expression différentielle et les relations entre des réseaux de gènes à partir de données d'expression longitudinales. Nous modélisons chaque profil d'expression génique comme une transformation aléatoire fonctionnelle de l'échelle, l'amplitude et la phase d'une courbe commune. Les inférences sur les paramètres d'amplitude gène-spécifique nous permettent d'étudier l'expression différentielle des gènes. Les inférences sur les mesures de similarité fonctionnelles basées sur des fonctions estimées de transformation du temps nous permettent d'étudier les réseaux de gènes tout en prenant en compte les caractéristiques des profils d'expression géniques. Nous appliquons notre méthode à des données simulées ainsi qu'à des données d'expression sur la progression du cancer de la prostate.

G. Xiao, C. Riley and A.B. Khodursky 805  
*Improved Detection of Differentially Expressed Genes Through Incorporation of Gene Locations*

La détermination des expressions différentielles dans les expériences de puces ADN fait habituellement l'hypothèse que le niveau d'expression d'un gène est indépendant du niveau d'expression des autres gènes. Cependant, de nombreuses études récentes ont montré que le niveau d'expression d'un gène tend à être similaire aux gènes situés dans la même région chromosomique et que les gènes différentiellement exprimés ont tendance à former des groupes d'intensité transcriptionnelle similaires le long du chromosome. Lorsqu'ils sont modélisés comme des séries spatiales à une dimension, les niveaux d'expression des gènes sur un même chromosome montrent souvent une corrélation spatiale significative, reflétant des motifs spatiaux dans la transcription. En modélisant ces corrélations spatiales, on peut obtenir de meilleurs estimateurs des niveaux de transcrits. Ici, nous démontrons l'existence de corrélations spatiales dans l'activité transcriptionnelle au sein du chromosome d'*Escherichia (E.) Coli*. dans plus de 50 conditions expérimentales. Sur la base de cette observation, nous proposons un modèle bayésien hiérarchique qui prend en compte l'information sur les gènes voisins pour améliorer l'estimation du niveau d'expression pour un gène donné et en conséquence la détection de gènes différentiellement exprimés. De plus, nous étendons le modèle pour prendre en compte la structure circulaire du chromosome d'*E. Coli* et la distance

intergénétique entre les gènes voisins. Les études de simulation et l'analyse des données réelles chez *E. coli* et la levure *Saccharomyces cerevisiae* montre que la méthode proposée surpasse la statistique *t* de SAM dans la détection de gènes différentiellement exprimés.

H-H. Muller, H. Schafer, J. Hebebrand and A. Scherag  
*Flexible Designs for Genomewide Association Studies*

815

Les études d'association à l'échelle du génome qui tentent de démêler l'étiologie génétique des traits complexes ont beaucoup attiré l'attention ces derniers temps. Ces études emploient fréquemment une stratégie de génotypage séquentielle : un large panel de marqueurs est examiné à l'aide d'un sous-échantillon de sujets, et les marqueurs les plus prometteurs sont ensuite génotypés chez les sujets restants. Dans cet article, nous introduisons une nouvelle méthode de ce type permettant entre autres de modifier les densités des marqueurs et les proportions d'échantillonnage tout en opérant un contrôle fort du taux d'erreur de type I FWER. Les pertes d'efficacité sont évitées en redistribuant les taux d'erreur de type I conditionnels des marqueurs abandonnés. Notre approche peut être combinée avec des plans d'expérience optimisant les coûts et offre une plus grande flexibilité que toutes les méthodes proposées précédemment. Entre autres, elle permet d'intégrer des critères biologiques à la procédure de sélection des marqueurs et offre la possibilité de modifier la taille de l'échantillon à tout moment pendant le déroulement du projet. Pour l'applicabilité pratique, nous développons un nouvel algorithme, que nous évaluons au moyen de simulations et illustrons à l'aide d'un jeu de données réel.

J-Y. Tzeng, D. Zhang, S-M Chang, D.C. Thomas and M. Davidian  
*Gene-Trait Similarity Regression for Multimarker-based Association Analysis*

822

Nous proposons une méthode de régression basée sur la similarité pour détecter l'association entre un trait et un génotype multi-marqueur. Notre approche régresse une mesure de similarité entre les traits de paires d'individus « non-apparentés » sur une mesure de similarité entre leurs haplotypes. La significativité est évaluée par un test de score dont nous décrivons la distribution asymptotique. La méthode proposée permet d'inclure des covariables, utilise une mesure de similarité indépendante des phases ce qui permet de s'affranchir de la reconstruction haplotypique et s'applique aussi bien aux traits quantitatifs que qualitatifs. Nous montrons également qu'il existe un lien étroit entre la régression de la similarité trait-génotype et l'analyse haplotypique avec effets aléatoires, alors que ces deux approches sont généralement considérées comme des outils de modélisation différents. Ce lien permet d'unifier les approches dites d'« haplotype sharing » avec celles de « variance component » et ainsi de décrire directement les propriétés des statistiques d'« haplotype sharing » même lorsque l'étude analytique du modèle de régression basé sur la similarité devient difficile.

L'utilisation d'ADN est une technique non invasive pour identifier les animaux utilisés dans les modèles de marquage-recapture pour lesquels un identifiant unique est nécessaire. Alors qu'il est possible d'obtenir de grande quantité d'ADN, la principale difficulté réside dans les erreurs de génotypage qui peuvent amener à des erreurs dans l'identification des individus. La principale source d'erreur est la perte allélique générée par l'échec de l'amplification de l'ADN à un ou plusieurs loci. Cela a pour conséquence de classer des individus hétérozygotes comme homozygotes à ces loci où un seul allèle a été réellement détecté. Si ces erreurs ne sont pas décelées et les génotypes analysés de manière naïve dans des modèles de marquage-recapture, cela peut induire une surestimation de la taille de la population. La méthode standard pour remédier à ce problème est d'exclure les échantillons de faible qualité d'ADN avec le risque d'éliminer une grande proportion des données. Il est préférable de garder ces échantillons de faible qualité puisqu'ils contiennent quand même une information partielle sur les génotypes. Au lieu d'essayer de minimiser les erreurs ou d'exclure des échantillons à problème, nous proposons de modéliser la perte allélique au sein de l'analyse. Nous décrivons une méthode basée sur le principe de l'"augmentation des données" qui nous permet de modéliser les données des échantillons pour lesquels il existe une ambiguïté dans la détermination des génotypes. La méthode est accompagnée d'une application à l'analyse de données du blaireau européen (*Meles meles*).

Des méthodes statistiques ont été développées et appliquées pour l'estimation de populations lorsqu'il est trop difficile ou trop coûteux de les dénombrer. Dans ces méthodes dénommées méthodes multi-listes dans les dispositifs épidémiologiques, les individus sont associés entre listes, et l'estimation de la taille de la population se fait en modélisant les comptages au sein de tables de contingence multidimensionnelles incomplètes (sur la base de schémas présence/absence sur les listes). Etant donné que les méthodes multi-listes supposent que les listes sont établies instantanément, il y a peu d'options possibles pour estimer la taille inconnue d'une population fermée en se basant sur des listes établies en continu (longitudinalement). Cependant dans les dispositifs épidémiologiques, les listes en temps continu sont un produit associé en routine du fonctionnement administratif. Les méthodes existantes se basent sur des analyses de durées entre événements suivies d'une seconde étape d'estimation de la taille de la population. Nous proposons une méthode alternative concernant le problème épidémiologique à deux aspects associant l'estimation de la taille de la population et l'identification chez les patients des facteurs liés à la durée (en jours) de l'intervalle entre visites dans une structure de soins. Un cadre bayésien est proposé pour modéliser les durées des intervalles puisque, pour de nombreux patients, les données sont éparpillées ; beaucoup de patients ne sont vus qu'une seule ou deux fois. La méthode proposée est appliquée aux données qui nous ont motivées pour illustrer son applicabilité. Puis une



étude par simulation explore les performances de l'estimateur sous différentes conditions. Enfin le petit chapitre de discussion propose des pistes pour poursuivre ce développement méthodologique dans le cadre de l'estimation de population en temps continu.

A. Cangelosi and M.B. Hooten

850

*Models for Bounded Systems with Continuous Dynamics*

Des modèles pour des processus naturels non linéaires, comme les dynamiques de population, ont recueilli beaucoup d'attention en mathématiques appliquées. Par exemple, la compétition entre espèces a été modélisée extensivement par des équations différentielles. Souvent, le scientifique a préféré modéliser les processus dynamiques sous-jacents (c'est-à-dire, les mécanismes théoriques) en temps continu. Il est à la fois d'un intérêt scientifique et mathématique d'implémenter de tels modèles dans un cadre statistique pour quantifier l'incertitude associée aux modèles en présence des observations. C'est-à-dire, à partir des observations discrètes provenant d'un processus sous-jacent continu, le processus non observé peut être formellement décrit tout en prenant en compte de multiples sources d'incertitude (par exemple, l'erreur de mesure, le choix du modèle, et le caractère stochastique inhérent aux paramètres du processus). En plus de la continuité, les processus naturels sont souvent bornés ; spécifiquement, ils tendent à avoir un support non négatif. Diverses techniques ont été implémentées pour prendre en compte les processus non négatifs, mais de telles techniques sont souvent limitées ou trop compromettantes. Cette étude offre une alternative aux pratiques habituelles de modélisation différentielle en utilisant une distribution gaussienne tronquée corrigée du biais pour modéliser les observations et le processus latent, les deux ayant un support borné. Les paramètres du processus continu sous-jacent sont caractérisés dans un contexte hiérarchique bayésien utilisant une approximation de Runge-Kutta du quatrième ordre.

M. Kaiser and P.C. Carageo

857

*Exploring Dependence with Data on Spatial Lattices*

L'application de modèles de champs markoviens aléatoires à des données spatiales sur treillis nécessite de faire des choix sur plusieurs aspects importants de la structure du modèle. Les techniques exploratoires existantes pour données spatiales ne guident pas directement l'utilisateur quant à ces choix. Nous présentons une statistique exploratoire directement liée à la structure des modèles de champs markoviens aléatoires, et basée sur des distributions conditionnelles de la famille exponentielle à un paramètre. Nous montrons que ce diagnostic exploratoire est assez informatif pour guider les choix nécessaires à la modélisation d'une structure spatiale incluant des termes de dépendance statistique. Dans cet article, nous développons la statistique de diagnostic et montrons sur des données simulées comment l'utiliser pour guider les choix de modélisation, puis nous réétudions une application publiée précédemment.

G. Yin and Y. Yuan

866

*A Latent Contingency Table Approach to Dose-finding for Combinations of Two Agents*

Des essais associant deux médicaments ont récemment beaucoup attiré l'attention en recherche oncologique. Il y a plusieurs raisons importantes pour associer différents agents dans un traitement : induire un effet synergique entre les médicaments, augmenter la dose intensité sans provoquer de toxicités croisées, et cibler différentes populations tumorales. Pour s'adapter à cette tendance croissante dans les essais cliniques, nous proposons un plan bayésien adaptatif pour la recherche de dose, basé sur des tables de contingence 2x2 latentes. Nous recherchons la dose maximale tolérée en combinant deux agents ; et nous adaptons continuellement les estimateurs a posteriori des paramètres inconnus associés aux probabilités marginales et du paramètre de corrélation basé sur les données de patients inclus successivement. En ré-ordonnant les probabilités de dose toxicité dans un espace à deux dimensions, nous attribuons à chaque cohorte de patients inclus la dose la plus appropriée pour la combinaison. Nous faisons de larges études de simulations pour examiner les caractéristiques opérationnelles de la méthode proposée sous différents scénarii pragmatiques. Enfin, nous appliquons notre procédure de recherche de dose à un essai clinique du M.D. Anderson Cancer Center

Y. Ji and B.N. Bekele

876

*Adaptive Randomization for Multi-arm Comparative Clinical Trials Based on Joint Efficacy Toxicity Outcomes*

Nous présentons ici une procédure de randomisation adaptative, l'adaptation se faisant au regard du critère d'évaluation, pour des essais cliniques comparatifs dont le critère d'évaluation principal est un critère combiné d'efficacité et de toxicité. Dans cette procédure, les probabilités d'allocation aux différents bras de traitement sont déséquilibrées de façon adaptative, de sorte que soient favorisés les traitements présentant les meilleurs résultats sur le critère d'évaluation, c'est-à-dire la plus grande efficacité et la plus faible toxicité. Ce type de procédure est avantageux pour les patients puisqu'en moyenne, ils sont randomisés en plus grand nombre dans les bras de traitement les meilleurs. En développant une extension du modèle bayésien approximé pour temps jusqu'à événement de Cheung et Thall (2002), nous modélisons le critère combiné d'efficacité et de toxicité que nous estimons a posteriori par une approche en variable latente. De ce fait, nous pouvons incorporer dans le modèle l'essentiel de l'information concernant les patients au suivi incomplet. C'est sur la base des probabilités a posteriori ainsi obtenues que nous construisons notre procédure de randomisation adaptative, favorisant les traitements avec les probabilités les plus élevées d'efficacité sans toxicité. Nous illustrons cette méthode sur un essai comparant trois traitements de la leucémie en termes de taux de rémission moléculaire et taux de toxicité à 52 semaines.

D. Follmann, M. Fay and M. Proschan

885

*Chop-lump Tests for Vaccine Trials*

Cet article propose de nouveaux tests pour comparer un vaccin à un placebo dans un essai randomisé de vaccination quand une faible proportion des participants est infectée. Une approche simple, cohérente avec une démarche en intention de traiter, est d'affecter un score  $W$  égal à zéro pour les non infectés ou égal à un critère de jugement post-infection  $X$ , positif, pour les infectés. On peut alors tester l'égalité des distributions

dissymétriques de  $W$  entre les deux groupes. Cette mesure du Fardeau de la Maladie (BOI - Burden Of Illness) a été introduite par Chang, Guess et Heyse (1994). Si les infections sont rares, le grand nombre de zéros dans les deux groupes dilue l'effet du vaccin et ce test peut avoir une puissance faible, en particulier si les  $X$  ne sont pas proches de zéro. La comparaison des  $X$  sur les seuls sous-groupes d'infectés n'est plus une comparaison de groupes randomisés, ce qui peut fausser les conclusions. Gilbert, Bosch et Hudgens (2003) et Hudgens, Hoering et Self (2003) ont introduit les tests de l'égalité des  $X$  sur le sous-groupe des sujets "condamnés" à être infectés dans toute randomisation. Cette démarche peut être plus puissante que celle du BOI, mais nécessite des conjectures non vérifiables. Nous suggérons une nouvelle version sur groupes amputés (chop-lump) de tests  $t$  et de Wilcoxon qui peut être plus puissante que le BOI dans certaines circonstances. Quand les effectifs sont égaux dans chaque groupe, le test ampute les deux groupes d'un nombre égal de zéros et réalise ensuite un test de comparaison des  $W$  qui sont alors pour la plupart positifs. Une évaluation de la distribution sous  $H_0$  est réalisée par permutations. Nous montrons que pour des alternatives locales le test de Wilcoxon "amputé" est plus puissant que le test usuel de Wilcoxon, à condition que les taux de contact soient les mêmes pour les groupes vaccinés et placebo. Nous mettons aussi en évidence l'influence cruciale de l'écart entre zéro et les  $X$  sur la puissance des tests  $t$ . Les tests amputés sont comparés aux tests antérieurs par une simulation effectuée sur des essais planifiés pour le VIH et la malaria. Une réanalyse d'un essai de phase III d'un vaccin pour le VIH est utilisée comme illustration de la méthode.

L. Tian, T. Cai and L.J. Wei

894

*Identifying Subjects Who Benefit from Additional Information for Better Prediction of the Outcome Variables*

Supposons que l'on soit intéressé par l'utilisation de nouveaux marqueurs biologiques ou cliniques pour améliorer la prédiction ou le diagnostic du devenir clinique du patient, en plus des marqueurs conventionnels. La valeur ajoutée des nouveaux marqueurs est typiquement mesurée en moyennant sur tous les patients de la population étudiée. Cependant, la mesure des nouveaux marqueurs étant coûteuse et invasive, une amélioration globale ne justifie pas la mesure des nouveaux marqueurs chez tous les patients. Une stratégie plus pratique est d'utiliser les marqueurs conventionnels du patient pour décider si les nouveaux marqueurs sont nécessaires ou non à l'amélioration de la prédiction de son état de santé. Dans cet article, nous proposons des procédures d'inférence pour les valeurs ajoutées des nouveaux marqueurs dans des sous-groupes variés de patients, classés par leurs marqueurs conventionnels. Le point qui en résulte et les estimations par intervalle peuvent être bien utiles pour ceux qui prennent des décisions médicales recherchant un équilibre entre la valeur diagnostique ou prédictive des nouveaux marqueurs et leurs coût et risque associés. Nos propositions sont justifiées théoriquement et illustrées empiriquement par deux exemples réels.

H-Y. Kim and M.G. Hudgens

903

*Three-Dimensional Array-Based Group Testing Algorithms*

Nous obtenons les caractéristiques opérationnelles d'algorithmes de test basés sur des tableaux ternaires, pour l'identification de « cas » en présence d'erreur de test. Les caractéristiques opérationnelles étudiées comprennent l'efficacité (ici le nombre de tests par spécimen) et les taux d'erreur (sensibilité, spécificité, valeurs prédictives négative et positive). Les méthodes sont illustrées en comparant les algorithmes proposés avec des algorithmes hiérarchiques et sur tableaux à deux dimensions, dans la détection de cas récents d'infection HIV en Caroline du Nord. Nos résultats montrent que les algorithmes basés sur des tableaux ternaires peuvent être plus efficaces et précis que les algorithmes précédemment proposés, dans des contextes avec erreur de test et faible prévalence.

G. Zhang and R. Little 911  
*Extensions of the Penalized Spline of Propensity Prediction Method of Imputation*

Little et An (2004) ont proposé une méthode d'imputation de valeurs manquantes pour prédire la propension par spline pénalisée (PSPP) qui conduit à une inférence robuste basée sur un modèle, sous l'hypothèse de mécanisme aléatoire. Le score de propension pour une variable manquante est estimé et un modèle de régression est ajusté incluant le logit estimé du spline du score de propension comme une covariable. La moyenne non-conditionnée prédite de la variable manquante a une propriété de double robustesse (DR) lorsque le modèle d'imputation est mal spécifié. Nous montrons qu'une version simplifiée de PSPP, qui ne centre pas les autres régresseurs avant de les inclure dans le modèle de prédiction, vérifie également la propriété DR. Nous proposons également deux extensions de PSPP, la PSPP stratifiée et la PSPP bivariée, qui étendent la propriété DR aux inférences sur les moyennes conditionnelles. Ces méthodes d'extension sont comparées à la méthode PSPP et à des alternatives simples par une étude de simulation et appliquées à une étude de perte de poids conduite par Kaiser Permanente.

H. Zhu, H. Zhou, J. Chen, Y. Li, M. Styner and J. Lieberman 919  
*Adjusted Exponentially Tilted Likelihood with Applications to Brain Morphology*

Dans cet article, nous développons une méthode nonparamétrique appelée vraisemblance ajustée inclinée exponentiellement, et nous l'appliquons à l'analyse de mesures morphométriques. Nous montrons que l'estimateur ajusté exponentiellement incliné possède les mêmes propriétés asymptotiques de premier ordre que l'estimateur exponentiellement incliné original. La statistique du quotient de vraisemblance ajustée inclinée exponentiellement est appliquée pour tester des hypothèses linéaires sur les paramètres inconnus, telles que l'association entre des mesures sur le cerveau (par exemple, les surfaces corticale et subcorticale) et des variables comme l'âge, le sexe, ou les gènes. Nos simulations montrent que la statistique du quotient de vraisemblance ajustée inclinée exponentiellement fonctionne aussi bien que le test t lorsque les données d'imagerie suivent une loi symétrique, et mieux lorsque la loi est asymétrique. A titre d'exemple, nous appliquons notre nouvelle méthode statistique à la détection de différences statistiquement significatives dans la morphologie de l'hippocampe entre deux types de schizophrénie et des patients sains.

*Forcing Function Diagnostics for Nonlinear Dynamics*

Cet article étudie le problème des diagnostics de modèles pour des systèmes décrits par des équations différentielles ordinaires (EDO). Nous proposons de modéliser l'écart au modèle par une correction variant dans le temps du membre de droite d'une équation différentielle proposée. Cette correction peut être décrite comme un ensemble de fonctions de forçage additives, estimé à partir des données. Représenter l'écart au modèle de cette façon nous permet d'étudier graphiquement les défauts d'adéquation du modèle et de suggérer des améliorations.

Nous en dérivons des tests d'écarts au modèle fondés sur des fonctions de forçage estimées. La construction de modèles dans des systèmes d'EDO partiellement observés est particulièrement difficile et nous étudions le problème de l'identification des fonctions de forçage dans ces systèmes. Les méthodes sont illustrées avec des exemples de neurosciences.

**Biometric Practice**

M. Rosenblum and M.J. van der Laan

937

*Using Regression Models to Analyze Randomized Trials: Asymptotically Valid Hypothesis Tests Despite Incorrectly Specified Models*

Les modèles de régression sont souvent utilisés pour déterminer les relations de causes à effets à partir de données collectées dans des essais ou des expériences randomisés. Cette pratique a fait l'objet d'un examen minutieux puisque les modèles usuels logistiques et de régression linéaire ne détectent pas souvent les véritables relations entre les variables. De plus des modèles spécifiés de façon incorrects conduisent potentiellement à des conclusions incorrectes

Dans cet article nous nous concentrons sur des tests d'hypothèse lorsque le traitement donné dans un essai randomisé a quelque effet sur la moyenne de résultats primaires dans des strates de variables de contrôle tel que l'âge, le sexe et l'état de santé. Notre premier souci est de s'assurer que de tels tests d'hypothèses présentent une erreur de type I correcte pour de grands échantillons. Notre principal résultat est que pour une importante classe de modèles de régression, modèle de régression standard basée sur des tests d'hypothèses (mais utilisant des estimateurs de variance robuste), on observe une erreur de type I correcte pour de grands échantillons même lorsque les modèles sont incorrectement spécifiés.

Au meilleur de notre connaissance la robustesse de tels modèles basés sur des tests d'hypothèse à partir de modèles incorrectement spécifiés étaient auparavant inconnue pour les modèles de régression de type poissonien et les modèles communément utilisés. Nos résultats ont des implications pratiques pour des modèles de base habituellement utilisés dans l'analyse d'essais randomisés.

*Combining Information from Cancer Registry and Medical Records Data to Improve Analyses of Adjuvant Cancer Therapies*

Les registres des cancers contiennent des informations intéressantes sur le volume des traitements adjuvants des patients atteints de cancer. Des études antérieures ont cependant montré que ces traitements étaient sous-déclarés dans les registres. L'utilisation directe de ces registres peut donc conduire à des résultats invalides. Nous proposons d'imputer la variable prise de traitement adjuvant à partir d'une autre source d'informations, par exemple un échantillon de validation de dossiers médicaux, avant d'analyser les imputations multiples selon l'approche de Yucel et Zalavsky (2005). Nous étendons leur modèle à la situation des thérapies multiples au moyen d'un modèle probit à effet aléatoire. Notre modèle prend en compte les associations entre les différents traitements tant au niveau de leur administration que de la probabilité de les déclarer. En outre nous prenons en compte la structure multi-niveaux des données de registre induite par la présence de malades traités au sein d'un même hôpital. Nous utilisons la méthode d'échantillonnage de Gibbs pour estimer le modèle et imputer la variable prise de traitement adjuvant. Nous appliquons la technique proposée sur des données du projet Quality of Cancer Care, dans lequel les malades de cancers colorectaux de stades II et III étaient éligibles à la chimiothérapie adjuvante et à l'irradiation.

*Optimal Bayesian Design for Patient Selection in a Clinical Study*

Un plan expérimental bayésien pour un essai clinique nécessite la spécification d'une fonction d'utilité qui modélise le but de l'essai, soit ici la sélection de patients pour un test diagnostique. Le meilleur échantillon de patients est alors sélectionné en maximisant l'utilité espérée. Ce travail d'optimisation se heurte à des difficultés dues à la dimension élevée de l'espace discret du plan d'expérience, et aussi à la complexité élevée de la formule d'utilité espérée. Une méthode pour un plan optimal, basée sur la simulation, peut être envisagée dans ce cas. De plus, nous développons deux algorithmes déterministes effectuant une recherche systématique sur l'espace de planification, pour évaluer les questions de calculs.

*Comparison of Hierarchical Bayesian Models for Over-dispersed Count Data using DIC and Bayes Factors*

Lorsque des données de comptage répliquées sont excessivement dispersées, il est de pratique courante d'incorporer cette variabilité extra-poissonnienne par l'inclusion de paramètres latents au niveau des observations. Par exemple, les modèles binomiale négative et Poisson-lognormal sont obtenus en utilisant des paramètres latents gamma et log-normal respectivement. Plusieurs publications récentes ont utilisé le critère de la déviance d'information (DIC) pour choisir entre ces deux modèles, la déviance étant définie en utilisant la vraisemblance de Poisson obtenue par conditionnement sur ces paramètres latents. Les résultats présentés ici montrent que cette utilisation du DIC est

inappropriée. En réalité on voit que le DIC se comporte très bien s'il est calculé en utilisant la vraisemblance marginalisée au niveau groupe en intégrant les paramètres latents pris au niveau des observations. Cette marginalisation au niveau groupe est explicite pour la binomiale négative, mais nécessite une intégration numérique pour le modèle Poisson-lognormal. De même le DIC se comporte très bien pour apprécier si l'inflation de zéro est nécessaire lorsque le calcul utilise la forme marginalisée de la vraisemblance à inflation de zéro. Dans le contexte de la comparaison de modèles hiérarchiques multi-niveaux, le DIC du niveau élevé a été obtenu à partir de la vraisemblance marginalisée ensuite par intégration sur les paramètres latents au niveau groupe, et d'autre part les densités marginales ont été calculées pour permettre d'obtenir les facteurs de Bayes. Les aspects calculatoires et l'interprétabilité de ces différentes mesures sont étudiés.

S. Luo, C.M. Crainiceanu, T.A. Louis and N. Catterjee

970

*Bayesian Inference for Smoking Cessation with a Latent Cure State*

Nous présentons une approche bayésienne pour modéliser les processus comportementaux dynamiques d'intoxication au tabac, quand la guérison n'est pas directement observée en raison des censures. Des probabilités spécifiques à chaque sujet modélisent les transitions stochastiques entre trois états comportementaux : fumeur, arrêt temporaire, arrêt définitif (état absorbant). Une distribution normale multivariée est appliquée aux effets aléatoires afin de prendre en compte les possibles corrélations entre les probabilités de transition spécifiques à chaque sujet. L'analyse est faite en utilisant une structure bayésienne à travers des simulations de Monte Carlo - Chaînes de Markov. Cette approche fournit diverses mesures des prédictions propres à chaque sujet ; ces dernières sont utiles pour des politiques de décisions, le développement interventionnel et l'évaluation. Des simulations sont réalisées pour valider notre méthodologie bayésienne et évaluer ses propriétés fréquentistes. Nos méthodes sont appliquées à une grande étude longitudinale de cohorte chez les fumeurs en Finlande (29133 sujets), l'étude de Prévention du Cancer du Poumon par Alpha-Tocophérol, Beta-Carotene (ATBC).

A. Kawaguchi, G.G. Koch and R. Ramaswamy

979

*Applications of Extensions of Bivariate Rank Sum Statistics to the Crossover Design to Compare Two Treatments Through Four Sequence Groups*

Cet article décrit des applications des extensions de statistiques bivariées de somme de rangs pour quatre groupes de séquence de deux traitements. Un essai clinique randomisé en ophtalmologie fournit la motivation de base pour la discussion. Le plan bilatéral pour cette étude a quatre groupes de séquence T:T, T:P, P:T et P:P respectivement, pour T le traitement à tester et P le placebo dans l'ordre correspondant pour les yeux de gauche à droite. Cet article décrit comment utiliser la moyenne des statistiques séparées des sommes de rang de Wilcoxon pour les yeux de gauche et de droite pour la comparaison globale de T et de P avec prise en compte de la corrélation entre les deux yeux. Une extension de ce critère avec une meilleure sensibilité aux différences potentielles entre T et P par réduction de la variance utilisable est discutée en fonction d'un modèle conceptuel avec des contraintes pour l'homogénéité interne des groupes mitoyens ayant le

même traitement et pour l'homogénéité externe des différences entre T et P. La qualité de l'ajustement pour ce modèle peut être évaluée avec les statistiques de test pour les contraintes correspondantes. Des études de simulation pour le modèle conceptuel confirment que la puissance est meilleure pour la statistique de test étendue que pour d'autres critères sans cette propriété. Les méthodes résumées ci-dessous sont illustrées par un essai clinique en ophtalmologie, mais ils sont applicables à d'autres situations avec un plan avec permutation des objets pour quatre groupes de séquence soit pour deux positions pour deux traitements simultanément pour un patient soit pour deux périodes successives des traitements prescrits pour un trouble récurrent. Cet article note également que les méthodes basées sur leur modèle conceptuel peuvent avoir une puissance non satisfaisante pour des écarts à ce modèle où la différence entre T et P à travers les groupes T:T et P:P groupes n'est pas similaire à celle à travers les groupes T:P et P:T, comme cela peut arriver lorsque T a un effet systématique dans un essai bilatéral. Pour cette situation, des statistiques de test plus robustes peuvent être trouvées, mais on peut voir que les plans en groupes parallèles avec seulement les groupes T:T et P:P peuvent être plus utiles que le plan bilatéral avec quatre groupes de séquence.