
Méthodes d'association génétique

Président: Angelo Canty (McMaster University)

ANTONIO CIAMPI, McGill University

Rôle des méthodes d'apprentissage statistique dans les études d'association génétique

Les études d'association pangénomiques (GWAS) emploient souvent des tests univariés pour identifier les associations entre SNP et caractères, même si la biologie suggère une interaction entre les gènes. Or, certaines méthodes puissantes, mais rarement utilisées, tirées de la littérature sur l'apprentissage statistique peuvent aider à identifier des relations multi-SNP et mieux comprendre les associations SNP-caractères. Nous démontrons la performance de telles méthodes sur des ensembles de SNP à environ 100ko d'une vingtaine de gènes candidats associés aux mesures de densité osseuse.

BINOD NEUPANE, McMaster University

Méta-analyse multivariée dans les études d'association génétique

Pour les méta-analyses, des approches multivariées peuvent être préférables aux approches univariées pour estimer des paramètres fondés sur des phénotypes corrélés. Dans les études d'association génétique, les données manquantes sont rares et les associations généralement très faible ou modestes, et peuvent être hétérogènes. Donc, on s'inquiète de détecter de telles associations et de les estimer de façon précise et sans biais. Nous utilisons des simulations pour étudier la performance de méthodes de méta-analyse multivariées. Nos résultats préliminaires suggèrent que sous certains scénarios, les techniques multivariées performant mieux en terme de puissance, de couverture et de biais. Nous présenterons des résultats détaillés de notre analyse comparative pour différents scénarios réalistes.

JULIA TALEBAN, Samuel Lunenfeld Research Institute

Réduction du biais par bootstrap dans une analyse d'association génétique du temps avant événement

Bien qu'il soit nécessaire de contrôler le taux de faux négatifs dans les études pangénomiques, la sélection de marqueurs génétiques (SNP) avec de petites valeurs-p d'association introduit un biais optimiste dans les valeurs estimées des paramètres. Motivés par une étude sur les complications du diabète, nous élargissons une méthode de rééchantillonnage bootstrap pour les caractères quantitatifs/binaires pour analyser le temps avant la néphropathie dans un modèle de risque proportionnel. Nous évaluons le biais relatif et absolu à l'aide de simulations. Pour les SNP de puissance faible à modérée, les estimations bootstrap sont plus précises que les estimations non corrigées, mais la méthode a tendance à surcorriger lorsque la puissance est élevée. Pour les faux positifs, toutefois, la réduction du biais réduit convenablement les estimations vers zéro.

ANDRIY DERKACH, University of Toronto

Combiner des tests linéaires et quadratiques pour des variantes rares produit un test robuste pour différents modèles génétiques

Les variantes rares jouent un rôle important sur les traits et maladies complexes chez l'humain. Bien que plusieurs tests d'association aient été proposés pour les variantes rares, il y a beaucoup de confusion au sujet du choix d'un bon test en pratique. Des évaluations récentes catégorisent les méthodes en deux classes : les statistiques linéaires, sensibles aux directions des alternatives, et les statistiques quadratiques omnibus. Toutefois, aucun type de test n'est uniformément meilleur que l'autre. Pour assurer la robustesse, nous considérons l'approche de la valeur p minimale et l'approche de Fisher pour combiner des valeurs p de tests linéaires et quadratiques. Des études analytiques et par simulation montrent que les deux méthodes sont robustes pour des modèles génétiques avec différents paramètres.

ZHIJIAN CHEN, Samuel Lunenfeld Research Institute

Plans stratifiés séquentiels à deux phases pour séquençage régional suite à une étude d'association pangénomique : une approche bayésienne

Dans les études de suivi ciblées, les chercheurs séquentent parfois une région génomique d'intérêt au niveau des paires de bases pour identifier des variantes causales. Nous étudions les plans bayésiens séquentiels à deux phases où certains sujets sont sélectionnés parmi les catégories de génotype d'un marqueur génétique, puis séquencés dans une région cible. À chaque point d'échantillonnage, nous appliquons la combinaison bayésienne pour tenir compte de l'incertitude génétique du modèle et une stratégie de répartition de la taille de l'échantillon adaptative pour améliorer l'efficacité d'estimation du paramètre d'association génétique. Nous présentons des résultats numériques pour divers caractères quantitatifs et binaires.

ZEYNEP BASKURT, University of Toronto

Approche de vraisemblance composite par le paradigme de preuves pour analyser l'association génétique dans les lignées

Royall (1997) a proposé le paradigme de preuves comme alternative aux paradigmes fréquentistes et bayésiens pour interpréter les données comme preuves. Celui-ci utilise le rapport de vraisemblance (RV) de deux hypothèses simples comme mesure objective de la solidité des preuves statistiques. Dans les études génétiques, les RV servent souvent à évaluer la preuve, mais celle-ci est difficile à établir en présence de structures familiales très complexes. Nous utilisons des vraisemblances composites pour construire des RV permettant l'analyse probante des familles. Nous expliquons comment protéger ces RV contre les erreurs de spécification du modèle et montrons qu'elles présentent de bonnes caractéristiques opérationnelles et sont compatibles avec les modèles concurrents.