

## Estimation des risques dans les formes familiales de cancer.

Y. Drouet<sup>a</sup>, V. Bonadona<sup>a,b</sup> and C. Lasset<sup>a,b</sup>

<sup>a</sup>Unité de Prévention et d'Epidémiologie Génétique  
Centre Léon Bérard  
28 rue laënnec - 69008 LYON

<sup>b</sup>Departement Biomaths-Santé, Equipe Epidémiologie et Santé Publique  
UMR CNRS 5558 Laboratoire de Biométrie et Biologie Évolutive  
Université Claude Bernard Lyon 1  
43 bd du 11 novembre 1918 - 69622 VILLEURBANNE cedex

e-mail : youenn.drouet@lyon.unicancer.fr

**Mots clefs** : Statistique en génétique, Données familiales, Risque de cancer.

L'unité de Prévention et d'Epidémiologie Génétique du Centre Léon Bérard a pour mission principale de prendre en charge les familles à risque héréditaire de cancer. Cette activité nécessite d'estimer les risques de cancers des personnes suivies, et contribue ainsi au développement de nouvelles méthodes en statistique génétique. Nous présentons l'utilisation de R dans ce contexte. En utilisant les packages `kinship2` [1] et `BayesMendel` [2], nous montrons comment il est possible de calculer avec R : (i) la probabilité d'être porteur d'une mutation sachant l'histoire familiale, et (ii) le risque de cancer d'un individu en fonction de son génotype.

Les limites de ces outils sont cependant rapidement atteintes. A ce jour en effet, il n'est pas possible d'estimer le risque individualisé, en l'absence de mutation identifiée, d'un individu ayant une histoire familiale évocatrice d'un risque héréditaire [3]. Cette situation concerne pourtant la majorité des individus suivis dans les centres d'oncogénétique. D'autre part, il n'existe aucune fonction R permettant le calcul et l'optimisation des vraisemblances conditionnelles, prospectives et rétrospectives dans le cas de données familiales [4].

En s'appuyant sur une partie du code en C++ développé par B. Bonaiti dans le progiciel `GENERISK`, nous avons développé un prototype de fonction R permettant le calcul récursif de la vraisemblance d'une famille par l'algorithme d'Elston-Stewart [5]. Cette fonction utilise du code écrit en FORTRAN, ce qui permet de réaliser ce calcul connu pour sa complexité dans un temps très court, environ 200 fois plus rapidement que le même code écrit en R. L'interface en R permet d'utiliser : (i) la fonction d'optimisation `nlnmb` pour obtenir les estimations des paramètres de la fonction de risque qui maximisent la vraisemblance (ii) le package `bootstrap` pour calculer les intervalles de confiance, et (iii) les possibilités graphiques avancées de R. A terme, cette fonction permettra d'estimer les risques de cancer même en l'absence de mutation identifiée, à l'instar de la méthode BOADICEA d'estimation du risque de cancer du sein [6]. Le calcul des vraisemblances conditionnelles, prospectives et rétrospectives, sera également disponible, de même que la "Genotype-restricted likelihood", seule méthode capable de fournir des estimations sans biais dans le cas de familles sélectionnées à partir de critères cliniques et de tests génétiques [7]. Cette nouvelle fonction constituera ainsi un outil d'aide à la décision pour le clinicien, et pourra également être utilisée librement par les chercheurs méthodologistes en statistique génétique.

## Références

- [1] Therneau T., Atkinson E., Sinnwell J., Matsumoto M., Schaid D. and McDonnell S. (2012). kinship2: Pedigree functions. R package version 1.3.7.
- [2] Parmigiani G., Chen S., Wang W., Katki H., Blackford A. Adapted from C code written by Omar Aguilar, Giovanni Parmigiani. (2012). BayesMendel: Determining Carrier Probabilities for Cancer Susceptibility Genes. R package version 2.0-7.
- [3] Drouet Y. (2012). Modélisation de la susceptibilité génétique non observée d'un individu à partir de son histoire familiale de cancer. Applications aux études d'identification pangénomiques et à l'estimation du risque de cancer dans le syndrome de Lynch. *Thèse de doctorat*, Université Claude Bernard Lyon 1.
- [4] P. Kraft et D. C. Thomas. (2000). Bias and efficiency in family-based genecharacterization studies : conditional, prospective, retrospective, and joint likelihoods. *Am J Hum Genet.* **66**:1119-1131.
- [5] Elston RC, Stewart J. (1971). A general model for the genetic analysis of pedigree data. *Hum Hered.* **21**:523-542.
- [6] Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, Risch HA, Eyfjord JE, Hopper JL, Southey MC, Olsson H, Johannsson O, Borg A, Pasini B, Radice P, Manoukian S, Eccles DM, Tang N, Olah E, Anton-Culver H, Warner E, Lubinski J, Gronwald J, Gorski B, Tryggvadottir L, Syrjakoski K, Kallioniemi OP, Eerola H, Nevanlinna H, Pharoah PD, Easton DF. (2008). The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer.* **98**:1457-66
- [7] Bonaïti B, Bonadona V, Perdry H, Andrieu N, Bonaïti-Pellié C. (2011) Estimating penetrance from multiple case families with predisposing mutations: extension of the 'genotype-restricted likelihood' (GRL) method. *Eur J Hum Genet.* **19**:173-9