

# Progress and open challenges in extremely high-dimensional medical outcome prediction

Kevin Bleakley

► **To cite this version:**

Kevin Bleakley. Progress and open challenges in extremely high-dimensional medical outcome prediction. Journées MAS et Journée en l'honneur de Jacques Neveu, Aug 2010, Talence, France. <inria-00509772>

**HAL Id: inria-00509772**

**<https://hal.inria.fr/inria-00509772>**

Submitted on 16 Aug 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journées MAS 2010, Bordeaux

Session : Classification

## **Progress and open challenges in extremely high-dimensional medical outcome prediction**

par **Kevin Bleakley**

Using biological data for medical decisions requires "extremely high" prediction accuracy; mistakes can lead to death. Very few current statistical methods are good enough to be used in life-threatening clinical decisions, e.g. choice of low vs high chemotherapy dose for breast cancer patients. Difficulties include (1) the above moral reason, (2) high-dimensionality of data ( $p \gg n$ ) and (3) the possibility that data does not contain enough information to construct a near-perfect classification rule. I will review the current state-of-the-art in high-dimensional biological decision-making, showing what statistical methods are being used, their success (or lack of), and suggest possible future research directions. In particular, I will describe Next Generation Sequencing approaches, their faster-than-exponential drop in cost, and implications for the next five years at the statistics/biology interface.

*Adresse :*

Kevin BLEAKLEY

PostDoc, Mines ParisTech/Institut Curie/INSERM U900

Institut Curie - Centre de recherche Biologie du développement - U900

26 rue d'Ulm, 75248 Paris cedex 05, France

E-mail : [kevbleakley@gmail.com](mailto:kevbleakley@gmail.com)

<<http://cbio.ensmp.fr/~kbleakley>>

Session : Classification