

Et si votre traitement médical était conçu rien que pour vous ?

Génomique, bioinformatique et statistiques multivariées au service de la médecine de précision

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques
tpineldahoui@gmail.com | github.com/Baudelaire12

Mercredi 25 Février 2026

Série LinkedIn STEM au Quotidien — Jour 3/40

Plan de la présentation

- 1 Mise en contexte
- 2 Concepts clés
- 3 Développement mathématique
- 4 Cas pratique
- 5 Conclusion
- 6 Références

Un même médicament, des effets très différents

Paradoxe médical

Le même antidépresseur fonctionne chez **seulement 50%** des patients. L'autre moitié peut ressentir des effets secondaires graves sans bénéfice thérapeutique (Cipriani et al., 2018).

- Le génome humain contient environ **3,2 milliards de paires de bases** et **20 000 à 25 000 gènes** (NHGRI, 2022).
- On estime à **4 à 5 millions** le nombre de variants génétiques (SNPs) qui différencient deux individus.
- La pharmacogénomique — science qui étudie comment nos gènes influencent notre réponse aux médicaments — pourrait concerner **90% des prescriptions** courantes (PharmGKB, 2023).

Question centrale

Comment les statistiques et la bioinformatique permettent-elles d'adapter chaque traitement au profil génétique unique d'un patient ?

Génomique et SNPs : l'alphabet de notre individualité

Définitions essentielles :

Génome : Ensemble complet de l'information génétique d'un organisme encodée dans l'ADN.

SNP (Single Nucleotide Polymorphism) :

Variation d'un seul nucléotide à une position donnée du génome, partagée par au moins 1% de la population.

GWAS (Genome-Wide Association Study) :

Étude statistique qui associe des variants génétiques à des traits ou maladies sur l'ensemble du génome.

Médecine de précision : Approche intégrant données génomiques, cliniques et environnementales pour personnaliser les soins.

Exemple : BRCA1/BRCA2

Les mutations des gènes BRCA1 et BRCA2 multiplient par **5 à 10** le risque de cancer du sein. Identifier ces variants permet une surveillance renforcée et des thérapies ciblées (NICE, 2019).

Test d'association génétique (modèle linéaire généralisé) :

$$y_i = \mu + \beta_g \cdot g_i + \sum_{k=1}^p \beta_k X_{ik} + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

où $g_i \in \{0, 1, 2\}$ est le génotype au SNP testé (0, 1 ou 2 allèles mineurs).

Problème du test multiple — Correction de Bonferroni :

$$\alpha_{\text{corrigé}} = \frac{\alpha}{M} = \frac{0.05}{10^6} = 5 \times 10^{-8}$$

pour $M \approx 10^6$ SNPs testés simultanément dans un GWAS.

Score de risque polygénique (PRS) :

$$\text{PRS}_i = \sum_{j=1}^M \hat{\beta}_j \cdot g_{ij}$$

où $\hat{\beta}_j$ est l'effet estimé du SNP j et g_{ij} le génotype de l'individu i au SNP j .

Clustering et classification en médecine de précision

Classification non supervisée des sous-types tumoraux :

Méthode **k-moyennes** appliquée aux profils d'expression génique :

$$\arg \min_{\mathcal{C}} \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2$$

Analyse en Composantes Principales (ACP) pour visualiser la variabilité génétique :

$$Z = XW, \quad \text{où } W = \text{matrice des vecteurs propres de } X^T X$$

Application concrète

Le projet **TCGA (The Cancer Genome Atlas)** a utilisé ces méthodes pour identifier **33 types de cancers** distincts et leurs sous-types moléculaires, permettant des thérapies ciblées (TCGA Research Network, 2013).

Cas pratique : leucémie myéloïde chronique (LMC)

Succès de la médecine de précision

La LMC est causée dans **95%** des cas par la translocation chromosomique **t(9;22)** produisant le gène de fusion *BCR-ABL1*.

Chronologie scientifique :

- ❶ **1960** : Identification du chromosome Philadelphie (Nowell & Hungerford)
- ❷ **1973** : Caractérisation de la translocation t(9;22) (Rowley, 1973)
- ❸ **2001** : Approbation de l'Imatinib (Gleevec) — inhibiteur ciblé de BCR-ABL1
- ❹ **Résultat** : Survie à 10 ans passée de **20%** à **85%** (Druker et al., 2006)

Contexte africain : Le projet **H3Africa** (Human Heredity and Health in Africa) constitue la plus grande base de données génomiques africaines pour personnaliser la médecine aux populations du continent (H3Africa Consortium, 2014).

Synthèse

La médecine de précision fusionne génomique, bioinformatique et statistiques pour transformer des traitements standardisés en thérapies individualisées. Le séquençage du génome humain, dont le coût est passé de 3 milliards USD en 2003 à moins de 600 USD en 2023, démocratise cette révolution.

Perspectives :

- Intégration des données multi-omiques (génomique, protéomique, métabolomique)
- Développement de PRS adaptés aux populations africaines — encore sous-représentées dans les GWAS
- IA pour la prédiction de la résistance aux antibiotiques à partir du génome bactérien
- Enjeux éthiques : protection des données génétiques et accès équitable aux thérapies

« Le génome est le manuel d'utilisation de votre corps. La bioinformatique est l'art de le lire. »

— Adapté de Francis Collins, *The Language of God*, 2006

Références bibliographiques

-  Cipriani, A., et al. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs. *The Lancet*, 391(10128), 1357–1366.
-  National Human Genome Research Institute. (2022). *The Human Genome Project*. NHGRI.
-  Druker, B. J., et al. (2006). Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *New England Journal of Medicine*, 355(23), 2408–2417.
-  TCGA Research Network. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics*, 45, 1113–1120.
-  H3Africa Consortium. (2014). Research capacity : Enabling the genomic revolution in Africa. *Science*, 344(6190), 1346–1348.
-  National Institute for Health and Care Excellence. (2019). *BRCA testing guidance*. NICE Guideline NG151.
-  Rowley, J. D. (1973). A new consistent chromosomal abnormality in chronic myelogenous leukaemia. *Nature*, 243, 290–293.

Merci pour votre attention

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques

`tpineldahoui@gmail.com`

`github.com/Baudelaire12`

LinkedIn : @Pinel Baudelaire DAHOUI

Série STEM au Quotidien — 40 jours