



HAL
open science

Utilisations actuelles et impact du Big Data en psychiatrie

Aisha Belhadi

► **To cite this version:**

Aisha Belhadi. Utilisations actuelles et impact du Big Data en psychiatrie. Médecine humaine et pathologie. 2018. dumas-02115073

HAL Id: dumas-02115073

<https://dumas.ccsd.cnrs.fr/dumas-02115073>

Submitted on 30 Apr 2019

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.

UNIVERSITÉ de CAEN NORMANDIE

FACULTÉ de MÉDECINE

Année 2017/2018

THÈSE POUR L'OBTENTION
DU GRADE DE DOCTEUR EN MÉDECINE

Présentée et soutenue publiquement le : 13 Décembre 2018

par

Mme Aisha BELHADI

Née le 10 Février 1985 à Sarcelles (*95 Val d'Oise*)

Utilisations actuelles et impact
du Big Data en Psychiatrie

Président : Madame le Professeur Perrine BRAZO

Membres : Madame le Professeur Lydia GUITTET-BAUD

Monsieur le Docteur Fabian GUENOLE

Monsieur le Docteur Pascal CRETE

Monsieur le Docteur Grégory BONNET

Directeur de thèse : Dr Pascal CRETE

Année Universitaire 2017 / 2018**Doyen**

Professeur Emmanuel TOUZÉ

Assesseurs

Professeur Paul MILLIEZ (pédagogie)

Professeur Guy LAUNOY (recherche)

Professeur Sonia DOLLFUS & Professeur Evelyne Emery (3^{ème} cycle)**Directrice administrative**

Madame Sarah CHEMTOB

PROFESSEURS DES UNIVERSITÉS - PRATICIENS HOSPITALIERS

M.	AOUBA Achille	Médecine interne
M.	AGOSTINI Denis	Biophysique et médecine nucléaire
M.	AIDE Nicolas	Biophysique et médecine nucléaire
M.	ALLOUCHE Stéphane	Biochimie et biologie moléculaire
M.	ALVES Arnaud	Chirurgie digestive
M.	BABIN Emmanuel	Oto-Rhino-Laryngologie
M.	BÉNATEAU Hervé	Chirurgie maxillo-faciale et stomatologie
M.	BENOIST Guillaume	Gynécologie - Obstétrique
M.	BERGER Ludovic	Chirurgie vasculaire
M.	BERGOT Emmanuel	Pneumologie
M.	BIBEAU Frédéric	Anatomie et cytologie pathologique
Mme	BRAZO Perrine	Psychiatrie d'adultes
M.	BROUARD Jacques	Pédiatrie
M.	BUSTANY Pierre	Pharmacologie
Mme	CHAPON Françoise	Histologie, Embryologie
Mme	CLIN-GODARD Bénédicte	Médecine et santé au travail
M.	COQUEREL Antoine	Pharmacologie
M.	DAO Manh Thông	Hépatologie-Gastro-Entérologie
M.	DAMAJ Ghandi Laurent	Hématologie
M.	DEFER Gilles	Neurologie

M. DELAMILLIEURE Pascal	Psychiatrie d'adultes
M. DENISE Pierre	Physiologie
M. DERLON Jean-Michel <small>Éméritat jusqu'au 31/08/2018</small>	Neurochirurgie
Mme DOLLFUS Sonia	Psychiatrie d'adultes
M. DREYFUS Michel	Gynécologie - Obstétrique
M. DU CHEYRON Damien	Réanimation médicale
M. DUHAMEL Jean-François <small>Éméritat jusqu'au 31/08/2018</small>	Pédiatrie
Mme ÉMERY Evelyne	Neurochirurgie
M. ESMAIL-BEYGUI Farzin	Cardiologie
Mme FAUVET Raffaèle	Gynécologie – Obstétrique
M. FISCHER Marc-Olivier	Anesthésiologie et réanimation
M. GÉRARD Jean-Louis	Anesthésiologie et réanimation
M. GUILLOIS Bernard	Pédiatrie
Mme GUITTET-BAUD Lydia	Epidémiologie, économie de la santé et prévention
M. HABRAND Jean-Louis	Cancérologie option Radiothérapie
M. HAMON Martial	Cardiologie
Mme HAMON Michèle	Radiologie et imagerie médicale
M. HANOUCZ Jean-Luc	Anesthésiologie et réanimation
M. HÉRON Jean-François <small>Éméritat jusqu'au 31/08/2018</small>	Cancérologie
M. HULET Christophe	Chirurgie orthopédique et traumatologique
M. HURAUULT de LIGNY Bruno <small>Éméritat jusqu'au 31/01/2020</small>	Néphrologie
M. ICARD Philippe	Chirurgie thoracique et cardio- vasculaire
M. JOIN-LAMBERT Olivier	Bactériologie - Virologie
Mme JOLY-LOBBEDEZ Florence	Cancérologie
Mme KOTTLER Marie-Laure	Biochimie et biologie moléculaire
M. LAUNOY Guy	Epidémiologie, économie de la santé et prévention
M. LE COUTOUR Xavier	Epidémiologie, économie de la santé et prévention
Mme LE MAUFF Brigitte	Immunologie
M. LEPORRIER Michel <small>Éméritat jusqu'au 31/08/2020</small>	Hématologie

M. LEROY François	Rééducation fonctionnelle
M. LOBBEDEV Thierry	Néphrologie
M. MANRIQUE Alain	Biophysique et médecine nucléaire
M. MARCÉLLI Christian	Rhumatologie
M. MARTINAUD Olivier	Neurologie
M. MAUREL Jean	Chirurgie générale
M. MILLIEZ Paul	Cardiologie
M. MOREAU Sylvain	Anatomie/Oto-Rhino-Laryngologie
M. MOUTEL Grégoire	Médecine légale et droit de la santé
M. NORMAND Hervé	Physiologie
M. PARIENTI Jean-Jacques communication	Biostatistiques, info. médicale et tech. de
M. PELAGE Jean-Pierre	Radiologie et imagerie médicale
Mme PIQUET Marie-Astrid	Nutrition
M. RAVASSE Philippe	Chirurgie infantile
M. REZNIK Yves	Endocrinologie
M. ROUPIE Eric	Thérapeutique
Mme THARIAT Juliette	Radiothérapie
M. TILLOU Xavier	Urologie
M. TOUZÉ Emmanuel	Neurologie
M. TROUSSARD Xavier	Hématologie
Mme VABRET Astrid	Bactériologie - Virologie
M. VERDON Renaud	Maladies infectieuses
Mme VERNEUIL Laurence	Dermatologie
M. VIADER Fausto	Neurologie
M. VIVIEN Denis	Biologie cellulaire
Mme ZALCMAN Emmanuèle	Anatomie et cytologie pathologique

PROFESSEUR DES UNIVERSITÉS

M. LUET Jacques Éméritat jusqu'au 31/08/2018	Médecine générale
--	-------------------

PROFESSEUR ASSOCIÉ DES UNIVERSITÉS A TEMPS PLEIN

M. VABRET François	Addictologie
--------------------	--------------

PROFESSEURS ASSOCIÉS DES UNIVERSITÉS A MI-TEMPS

M. de la SAYETTE Vincent	Neurologie
Mme DOMPMARTIN-BLANCHÈRE Anne	Dermatologie
Mme LESCURE Pascale	Gériatrie et biologie du vieillissement
M. SABATIER Rémi	Cardiologie

PRCE

Mme LELEU Solveig	Anglais
--------------------------	---------

Année Universitaire 2017 / 2018

Doyen

Professeur Emmanuel TOUZÉ

Assesseurs

Professeur Paul MILLIEZ (pédagogie)

Professeur Guy LAUNOY

Professeur Sonia DOLLFUS & Professeur Evelyne Emery (3^{ème} cycle)

Directrice administrative

Madame Sarah CHEMTOB

MAITRES DE CONFERENCES DES UNIVERSITÉS - PRATICIENS

HOSPITALIERS

M.	ALEXANDRE Joachim	Pharmacologie clinique
Mme	BENHAÏM Annie	Biologie cellulaire
M.	BESNARD Stéphane	Physiologie
Mme	BONHOMME Julie	Parasitologie et mycologie
M.	BOUVIER Nicolas	Néphrologie
M.	COULBAULT Laurent	Biochimie et Biologie moléculaire
M.	CREVEUIL Christian	Biostatistiques, info. médicale et tech. de communication
Mme	DEBRUYNE Danièle <small>Éméritat jusqu'au 31/08/2019</small>	Pharmacologie fondamentale
Mme	DERLON-BOREL Annie <small>Éméritat jusqu'au 31/08/2020</small>	Hématologie
Mme	DINA Julia	Bactériologie - Virologie
Mme	DUPONT Claire	Pédiatrie
M.	ÉTARD Olivier	Physiologie
M.	GABEREL Thomas	Neurochirurgie
M.	GRUCHY Nicolas	Génétique
M.	GUÉNOLÉ Fabian <small>sera en MAD à Nice jusqu'au 31/08/18</small>	Pédopsychiatrie
M.	HITIER Martin	Anatomie - ORL Chirurgie Cervico-faciale
M.	LANDEMORE Gérard <small>sera en retraite à partir du 01/01/18</small>	Histologie, embryologie, cytogénétique
M.	LEGALLOIS Damien	Cardiologie
Mme	LELONG-BOULOUARD Véronique	Pharmacologie fondamentale
Mme	LEPORRIER Nathalie <small>Éméritat jusqu'au 31/10/2017</small>	Génétique

Mme LEVALLET Guénaëlle	Cytologie et Histologie
M. LUBRANO Jean	Chirurgie générale
M. MITTRE Hervé	Biologie cellulaire
M. REPESSÉ Yohann	Hématologie
M. SESBOÜÉ Bruno	Physiologie
M. TOUTIRAIS Olivier	Immunologie

MAITRES DE CONFERENCES ASSOCIÉS DES UNIVERSITÉS A MI-TEMPS

Mme ABBATE-LERAY Pascale	Médecine générale
M. COUETTE Pierre-André	Médecine générale
M. GRUJARD Philippe	Médecine générale
M. LE BAS François	Médecine générale
M. SAINMONT Nicolas	Médecine générale

Remerciements

À Rahma qui m'a faite mère

À Mlle G. qui m'a faite soignante

À Reda qui me fait questionner

Je suis profondément reconnaissante à mon directeur de thèse, Dr Pascal Créte, d'avoir accepté cette tâche et d'avoir pu partager son temps et son savoir, toujours avec patience et générosité. Plus que ça, je le remercie d'avoir été un enseignant, un directeur de stage, un collègue d'une grande humanité. Je ne pouvais espérer meilleur guide pour me lancer dans la découverte du mouvement de la psychothérapie institutionnelle (et digérer un peu de Lacan en deux heures de séminaire) !

Je remercie le Pr Perrine Brazo d'avoir accepté de présider ce jury. Je lui suis très reconnaissante d'avoir toujours facilité mon parcours d'interne et d'avoir été accessible, à la fois en tant que clinicienne et en tant qu'universitaire. Un grand merci au Dr Fabian Guérolé de s'être rendu disponible pour cette soutenance, mais aussi d'avoir émaillé mon internat de rencontres cliniques et d'enseignements très enrichissants.

Je suis très reconnaissante au Pr Lydia Guittet-Baud et au Dr Grégory Bonnet d'avoir accepté cette invitation qui leur est tombée des nues. J'espère que la lecture de ce travail éveillera et soutiendra leur intérêt jusqu'à la dernière page.

Je tiens à remercier tous les soignants et collègues que j'ai côtoyés à travers mes années d'internat, qui ont chacun modelé ce sens clinique et cette pratique que j'essayais de faire miens : l'UH1, la Maison des Adolescents et les foyers de l'ACSEA, le CRA et le CHR, le secteur Côte Fleurie à l'EPSM, l'Accueil de Jour pour Enfants et le CCTEF, le Foyer Léone Richet et l'Appui, et enfin le CMPEA. Et Mme Curto de la scolarité : merci !

À ma famille qui m'a soutenue et qui a surtout dû me supporter : it's done ! Je ne pourrai jamais prendre la mesure de tes sacrifices, Mama, encore moins pouvoir t'en remercier. And Baba, you reap what you sow : from playing around with 5-inch floppies and DOS commands, to this... Biology may not have been my destiny, but code certainly did call to me.

Enfin, je remercie ceux qui, pour notre fonction de soignants, sont notre raison d'être, et parfois aussi notre raison : tous les malades que j'ai rencontrés et que je rencontrerai. Leurs leçons sont inestimables. J'espère me montrer digne de leur confiance et ne jamais perdre de vue leur humanité et la mienne. Ils sauront me le rappeler.

Abréviations

A-CHESS	Addiction-Comprehensive Health Enhancement Support System
ADN	acide désoxyribonucléique
AMM	autorisation de mise sur le marché
ANN	réseau de neurones artificiels ou artificial neural network
ANOVA	analysis of variance
ANSM	Agence Nationale de Sécurité du Médicament et des Produits de Santé
ARN	acide ribonucléique
ARNm	ARN messenger
ASIP Santé	Agence Française de la Santé Numérique
AUC	<i>area under curve</i> , aire sous la courbe ROC
AVC	accident vasculaire cérébral
B-SNIP	Bipolar-Schizophrenia Network on Intermediate Phenotypes
BACS	Brief Assessment of Cognition in Schizophrenia
BOINC	Berkeley Open Infrastructure for Network Computing
CAARMS	Comprehensive Assessment of At-Risk Mental State
CANTAB	Cambridge Neurocognitive Test Automated Battery
CDC	Centers for Disease Control and Prevention
CIM	Classification Statistique Internationale des Maladies et des Problèmes de Santé Connexes
CNIL	Commission Nationale de l'Informatique et des Libertés
COMED	Combining Medications to Enhance Depression Outcomes
CSBS DP	Communication and Symbolic Behavior Scales Developmental Profile
DASS-21	Depression, Anxiety and Stress Scale version 21
DCM	dynamic causal modeling
DSM-V	Diagnostic and Statistical Manual of Mental Disorders, 5th edition
EBM	evidence-based medicine
ECG	électrocardiogramme
ECT	électroconvulsivothérapie
EEG	électroencéphalogramme
FDA	Food and Drug Administration
GABA	gamma-aminobutyric acid
GAFAM ou GAFAM	Google, Apple, Facebook, Amazon, Microsoft
GAM	modèles additifs généralisés ou <i>generalized additive models</i>
GENDEP	Genome-based Therapeutic Drugs for Depression
GPGPU	general-purpose computing on graphics processing units
GPRS	genomic profile risk score
GWAS	genome-wide association study
HAS	Haute Autorité de Santé
IA	intelligence artificielle ou <i>artificial intelligence</i> (AI)
IBIS	Infant Brain Imaging Study
IdO	Internet des Objets ou <i>Internet of Things</i> (IoT)
Inserm	Institut National de la Santé et de la Recherche Médicale
IRM	imagerie par résonance magnétique (IRMf : IRM fonctionnelle)
KNN	méthode des <i>k</i> plus proches voisins ou <i>k-nearest neighbor</i>
LDA	analyse discriminante linéaire ou <i>linear discriminant analysis</i>
LOOCV	leave-one-out cross-validation
LSA	<i>latent semantic analysis</i> ou analyse sémantique latente
M-CHAT	Modified Checklist for Autism in Toddlers
MDPH	Maisons Départementales des Personnes Handicapées

MSE	<i>mean squared error</i> ou erreur quadratique moyenne
NAPLS	North American Prodrome Longitudinal Study
NCBI	National Center for Biotechnology Information
NHS	National Health Service
NIH	National Institutes of Health
NIMH	National Institute of Mental Health
NLM	National Library of Medicine
OCR	reconnaissance optique de caractères ou <i>optical character recognition</i>
PANSS	Positive and Negative Syndrome Scale
PGC	Psychiatric Genome-Wide Association Consortium
PHQ-9	Patient Health Questionnaire
PMI	Precision Medicine Initiative
PMSI	programme de médicalisation du système d'information
PRIME	Personalized Real-time Intervention for Motivational Enhancement
QDA	analyse discriminante quadratique ou <i>quadratic discriminant analysis</i>
RDoC	Research Domain Criteria
RGPD	Règlement Général sur la Protection des Données
ROC curve	<i>receiver operating characteristic curve</i> , courbe de la sensibilité en fonction de 1-spécificité)
RSS	<i>residual sum of squares</i> ou somme résiduelle des carrés
rTMS	repetitive transcranial magnetic stimulation
SAIL Databank	Secure Anonymised Information Linkage Databank
SMS	Short Messaging Service
SNDS	Système National des Données de Santé
SNPs	single-nucleotide polymorphisms
SPANR	splicing-based analysis of variants
SROC curve	summary receiver operating characteristic curve
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
SVM	machine à vecteurs de support ou <i>support vector machine</i>
TALN ou NLP	traitement automatique du langage naturel ou <i>natural-language processing</i>
TCC	thérapie cognitivo-comportementale
TDAH	trouble déficitaire de l'attention avec hyperactivité
TOC	trouble obsessionnel-compulsif
UE	Union Européenne
VBM	voxel-based morphometry

Tableaux et figures

Figure 1 : Santé publique et population mondiale

Figure 2 : *Information Commons* et *Knowledge Network*

Figure 3 : Tendances des recherches sur Google pour « big data »

Figure 4 : Réseau de concepts définissant le Big Data

Figure 5 : Capacité de traitement des ordinateurs personnels de 1975 à 2010

Figure 6 : Infographie résumant les points forts du RGPD

Figure 7 : Écosystème du Big Data

Figure 8 : Surveillance Sentinelle des suicides en France

Figure 9 : Nombre n d'échantillons par rapport au nombre de p paramètres

Figure 10 : Compromis biais-variance et différence entre erreur d'entraînement et erreur test

Figure 11 : Données éparses en haute-dimension (sparsity)

Tableau 1 : Résumé des techniques d'apprentissage statistique selon les utilisations

Figure 12 : Réseau de neurones artificiel

Figure 13 : Étapes clés de l'apprentissage automatique

Tableau 2 : Initiatives de partage de données d'imagerie et consortiums de recherche, pour les troubles neurologiques et psychiatriques

Figure 14 : Représentation quantitative des caractéristiques du discours de patients sous forme de réseaux de mots

Figure 15 : Protocole d'apprentissage de l'outil SPANR

Figure 16 : Courbe sROC des résultats de méta-analyse des études de classification du trouble bipolaire en imagerie

Figure 17 : Distribution de patients selon les dimensions *anxiety-related* et *anhedonia-related* des caractéristiques de connectivité fonctionnelle

Figure 18 : Performance de différents modèles de classification de pathologies neurologiques et psychiatriques sur les données d'entraînement (validation interne)

Figure 19 : Distribution de patients en 6 clusters selon les composantes *anhedonia*, *anxious arousal* et *tension*

Figure 20 : Protocole de traitement des données et d'entraînement du modèle de prédiction

Figure 21 : Résultats principaux de performance sur les données d'entraînement et cohortes test

Figure 22 : Captures d'écran de l'application PRIME

Figure 23 : Investissements récents dans des start-ups en technologies de la santé mentale

Figure 24 : « Move Fast and Break Things »

Sommaire

Introduction.....	1	c. Génétique.....	87
I – Caractérisation du Big Data et des sciences des données.....	8	d. Biomarqueurs sériques.....	92
1. Cadre historique et propriétés principales	8	e. Approches multimodales.....	94
a. Pourquoi parle-t-on de Big Data ?.....	8	3. Prédire.....	100
b. Comment en est-on arrivé là ?.....	10	a. Prévention primaire et secondaire : prédire la survenue et le cours de la maladie.....	101
c. Cadre légal.....	17	b. Optimisation du traitement : prédire la réponse à une intervention.....	112
2. Spécificités des données en santé et en recherche biomédicale.....	22	4. Suivre.....	119
a. Données génétiques et d'imagerie : haute précision et haute dimensionalité	23	III – Outil révolutionnaire ou dangereux mirage ?.....	130
b. Dossiers médicaux informatisés et publications : le défi du texte libre et de la fragmentation.....	25	1. Cultures et pratiques.....	132
c. Données collectées par smartphone ou appareils connectés.....	27	2. Subversion de la raison.....	140
d. Données de santé publique.....	30	3. Être atomisé.....	152
3. Outils d'exploitation des données massives.....	32	Conclusion.....	161
a. Infrastructure de stockage et de traitement.....	32	Bibliographie.....	163
b. Traitement algorithmique et techniques de visualisation des données.....	35	Annexe 1 : Synthèse des études en revue... ..	187
Extraction de données.....	36	Annexe 2 : Matrice RDoC.....	192
Visualisation de données.....	37	Annexe 3 : Architectures de réseaux de neurones artificiels pour apprentissage profond	195
Analyse de données.....	39		
c. Initiatives de partage des données....	56		
II – Utilisations du Big Data en psychiatrie.....	59		
1. Découvrir.....	60		
2. Classifier.....	73		
a. Dossiers médicaux électroniques.....	74		
b. Imagerie.....	78		

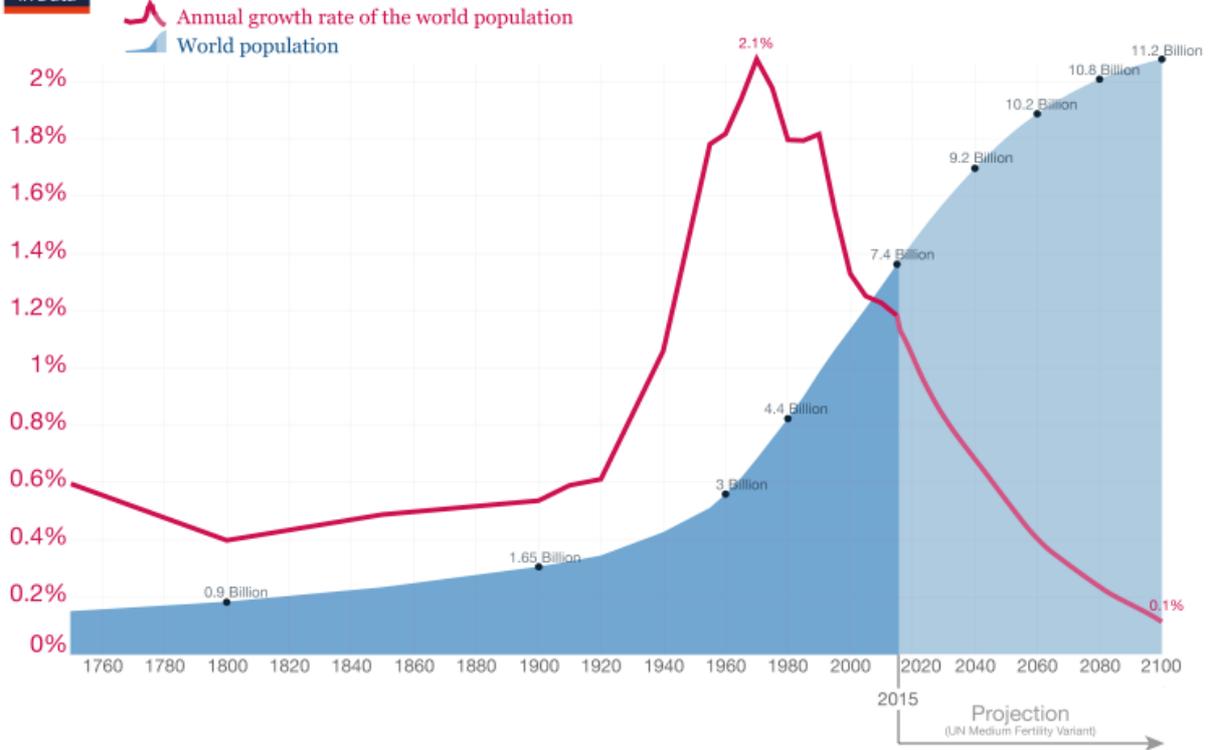
Introduction

L'augmentation de l'espérance de vie et l'explosion démographique du 20^{ème} siècle (« Espérance de vie - Mortalité–Tableaux de l'Économie Française | Insee », s. d.; United Nations, 1999), peuvent être en grande partie attribuées aux avancées en santé publique et en médecine. Les mesures de prévention, telles que la vaccination, la sécurité sanitaire et alimentaire, l'hygiène, la lutte contre la malnutrition et contre les risques professionnels et environnementaux, ont permis le recul du risque infectieux et de la mortalité infantile (CDC, s. d.). Ces grandes initiatives publiques, dont certaines à l'échelle globale, comme les campagnes de vaccination (World Health Organization, 2009; « World Population Growth », s. d.), Figure 1) faisaient suite aux découvertes scientifiques du 19^{ème} siècle en physiologie et en infectiologie, et à l'apparition d'outils mathématiques, les statistiques, permettant d'organiser et d'interpréter des données afin d'en extraire une information nouvelle (Khoury & Ioannidis, 2014). La propagation à travers le monde du modèle de l'état-nation, et son organisation bureaucratique, a vraisemblablement permis un niveau de contrôle administratif des populations et des territoires facilitant la mise en place d'une politique de santé publique nationale et internationale.

De manière illustrative, le Center for Disease Control (CDC, agence américaine de sécurité sanitaire et de santé publique) identifie dix grands succès de santé publique au cours du 20^{ème} siècle (CDC, 2017) :

- Immunizations – Vaccinations
- Motor-Vehicle Safety – Sécurité routière
- Workplace Safety – Sécurité sur le lieu de travail
- Control of Infectious Diseases – Contrôle des maladies infectieuses
- Declines in Deaths from Heart Disease and Stroke – Diminution des décès dus aux maladies cardiovasculaires et aux AVC (accidents vasculaires cérébraux)
- Safer and Healthier Foods – Alimentation plus sûre et saine
- Healthier Mothers and Babies – Amélioration de la santé materno-infantile
- Family Planning – Planning familial
- Fluoridation of Drinking Water – Addition de fluor à l'eau potable
- Tobacco as a Health Hazard – Identification du tabac comme risque sanitaire

World population growth, 1750-2100



Data sources: Up to 2015 OurWorldInData series based on UN and HYDE. Projections for 2015 to 2100: UN Population Division (2015) – Medium Variant. The data visualization is taken from OurWorldInData.org. There you find the raw data and more visualizations on this topic.

Licensed under CC-BY-SA by the author Max Roser.



Figure 1 : Santé publique et population mondiale

Explosion démographique à partir du début du 20^{ème} siècle (haut, sous licence Creative Commons « World Population Growth », s. d.) ; Poster pour campagne de vaccination (gauche, World Health Organization, 2009)

La plupart de ces exemples correspondent à des facteurs de risque à effet important sur une ou plusieurs pathologies, touchant une grande partie de la population, et sous-tendus par une explication biologique ou physiologique en apparence simple et

plausible. Ces facteurs, dont plusieurs sont des facteurs environnementaux, sont facilement accessibles à des interventions globales à visée préventive : vaccinations, addition d'iode ou de fluor à l'alimentation, obligation de port des ceintures et casques de sécurité, asepsie en chirurgie etc. L'épidémiologie classique trouve sa puissance dans l'homogénéité.

Les études épidémiologiques cas-témoins, les statistiques nationales et la meilleure connaissance de la physiologie avaient permis d'identifier les risques infectieux, accidentels et nutritionnels. Par la suite, la constitution de cohortes longitudinales de longue durée, comme la Framingham Heart Study, a révélé les facteurs de risque cardiovasculaires, cause principale de morbi-mortalité à partir des années 50 (« History of the Framingham Heart Study », s. d.). Cette approche illustre l'importance de suivis individuels prolongés pour la découverte de mécanismes physiopathologiques.

Le premier séquençage du génome humain (Human Genome Project, (« NIH Fact Sheets - Human Genome Project », s. d.), lancé par les National Institutes of Health (NIH, institution américaine de recherche en santé) et d'autres acteurs, à l'aube du 21^{ème} siècle promettait une nouvelle vague de découvertes fondamentales qui révolutionneraient la compréhension, le diagnostic et le traitement des maladies. En plus des facteurs de risque externes, on dévoilerait ainsi les facteurs de risque génétiques, « internes », et on avancerait les connaissances sur le fonctionnement cellulaire et moléculaire. L'attente, notamment au niveau du grand public, était que ce projet livrerait des cibles thérapeutiques directement exploitables. La réalité était plutôt une masse d'information difficile à interpréter due à la polygénéricité et la pleïotropie des maladies, dévoilant des réseaux d'interactions complexes et redondants, des gènes de susceptibilité à effet subtil, ou des mutations clairement pathologiques mais rares (Stern, Linker, Vadodaria, Marchetto, & Gage, 2018).

Malgré les progrès remarquables de l'approche historique de santé publique et du virage génétique et biologique récent, nous sommes donc loin d'une compréhension satisfaisante des causes et des mécanismes de morbidité et de mortalité, surtout pour les maladies chroniques qui dominent maintenant le paysage pathologique (Vos et al., 2017). Au contraire, comme l'illustrent si bien le problème de résistances aux antibiotiques et les effets variables du microbiote sur l'organisme, nous nous retrouvons face à la complexité de systèmes évolutifs, dynamiques et interdépendants, qui fait émerger des scénarios difficiles à prévoir et à gérer.

L'une des raisons pour cette complexité est la variété, la variabilité des organismes biologiques et leur « dégénérescence » (degeneracy, définie par la capacité qu'ont des éléments de structure différente à accomplir les mêmes fonctions ou arriver au même résultat sous certaines conditions, en pouvant par ailleurs avoir d'autres fonctions dans d'autres conditions ; Edelman & Gally, 2001 ; Marom, 2009). Ces propriétés, cette flexibilité, qui font la robustesse d'entités issues de l'évolution par sélection naturelle, posent un défi à notre capacité de générer des connaissances et des modèles explicatifs à partir d'échantillons statiques limités, et de corrélations linéaires. Un pessimiste se demanderait même si cela ne remet pas en question notre capacité d'abstraction et de conceptualisation du vivant.

Le problème est prégnant en psychiatrie où la question de nosologie, la catégorisation des maladies mentales, donc leur abstraction, suscite de vifs débats. La dernière édition du manuel statistique et diagnostique (Diagnostic and Statistical Manual of Mental Disorders, DSM-V, American Psychiatric Association, 2013), a été critiquée par des psychiatres de tout bord (Frances, 2013). Le National Institute of Mental Health (NIMH, institution américaine de recherche en santé mentale), cherchant un autre angle conceptuel, a développé les Research Domain Criteria (RdoC, Annexe 2, Cuthbert & Insel, 2013 ; Insel et al., 2010), conceptualisés comme une matrice confrontant des mécanismes fondamentaux de fonctionnement cognitif à différents niveaux d'analyse, génétique, moléculaire, cellulaire, physiologique et comportementale, potentiellement plus adaptée à la recherche en santé mentale. Le chevauchement de la susceptibilité génétique et des données d'imagerie entre différentes pathologies psychiatriques (par exemple, entre schizophrénie et trouble bipolaire) interroge la validité biologique de ces catégories diagnostiques. Parallèlement, la limite de leur intérêt clinique est apparente dans la pratique psychiatrique quotidienne, notamment auprès de populations particulières (enfants, migrants, contextes de précarité, pour n'en citer que quelques unes).

En effet, ces groupes humains exposent – de manière peut être plus évidente et immédiate que dans la population générale – les failles des tentatives de généralisation et de standardisation en médecine. Par leur fluidité, leur instabilité et leur place à l'intersection de différentes problématiques médicales et sociales, ces populations se prêtent moins facilement aux techniques habituelles d'étude, et n'entrent pas tout à fait dans les « cases », dont elles soulignent ainsi les défaillances fondamentales.

Avec un cadre de compréhension insuffisant, comment mieux soigner, et surtout comment atteindre le graâl de la médecine, mieux prévenir les maladies au niveau collectif

et individuel ? Il apparaîtrait donc que les statistiques à l'échelle d'une population ou d'un échantillon – traitant de moyennes, de médianes et de distributions normales – n'apportent plus une réponse adéquate aux besoins d'efficacité et d'optimisation de systèmes de santé qui font face à un décalage entre la demande et les capacités de soins. Ces techniques gèrent mal la variabilité inter- et intra-groupe, et ne rendent pas assez compte de la complexité des questions de santé.

La proposition prédominante pour pallier à ce manque de granularité est d'aller vers une médecine de plus en plus renseignée par des données précises, exhaustives et surtout personnalisées. Il ne s'agit plus seulement de recourir à de larges cohortes avec un recueil de données à intervalles de temps réguliers, et un traitement global de ces données pour obtenir des tendances au niveau de la population. Au contraire, l'intégration verticale d'un grand nombre de données de sources diverses concernant une même personne, et l'agrégation de toutes ces données pour un grand nombre de personnes, serait la clé d'une pratique médicale plus efficace, répondant mieux aux particularités individuelles des patients, davantage basée sur un socle de connaissances scientifiques et nourrissant en retour ce corpus de connaissances.

C'est du moins le postulat qu'avance, entre autres, la Precision Medicine Initiative (PMI) lancée par le NIH et le National Research Council (« National Institutes of Health (NIH) - All Of Us | National Institutes of Health (NIH) », s. d.; National Research Council (U.S.), 2011). Le but de ce projet est de créer une nouvelle nosographie qui dépasserait la sémiologie classique et intégrerait de manière continue et dynamique les informations moléculaires, biologiques et physiopathologiques les plus récentes, avec un niveau de précision à l'individu-près. L'infrastructure sous-tendant cette nouvelle nosographie serait une « commune d'information » (*Information Commons*) et un « réseau de connaissance » (*Knowledge Network*), qui seraient construits au cours de la pratique clinique quotidienne, en rendant systématiques la collecte d'échantillons individuels pour séquençage génétique, l'informatisation des dossiers médicaux incluant toutes les observations cliniques et examens complémentaires, et leur centralisation dans des bases de données accessibles aux cliniciens et aux chercheurs (Figure 2).

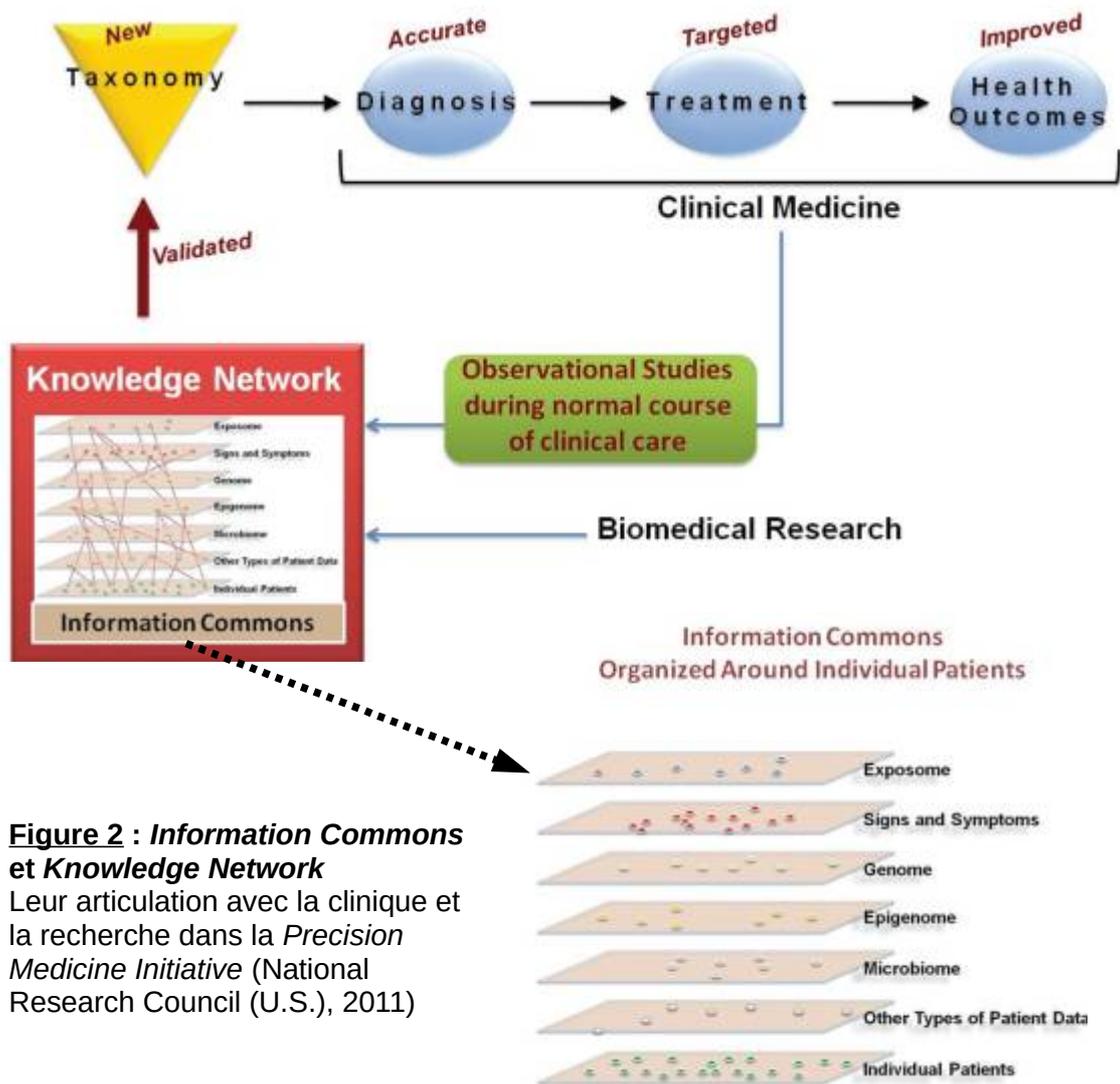


Figure 2 : Information Commons et Knowledge Network
 Leur articulation avec la clinique et la recherche dans la *Precision Medicine Initiative* (National Research Council (U.S.), 2011)

Cette vision n'est pas sans rappeler la réalité actuelle de nos profils numériques, qui concentrent, auprès de quelques géants du numérique, une vaste quantité de nos données intimes, exploitées pour en tirer de nouvelles informations et alimentant des objectifs commerciaux. Ce cadre, rendu possible par des innovations en matériel, logiciel, programmation et connectivité, est souvent connu sous l'appellation de Big Data (données massives) dans les publications scientifiques et techniques, et dans la presse (Lohr, 2013; Manyika et al., 2011; Passos, Mwangi, & Kapczynski, 2016). Ce terme simple, vague, sinistre pour les uns, avant-garde pour les autres, a trait à la génération massive de données numériques depuis le développement et la démocratisation des technologies numériques (« Big data | CNIL », s. d.; Helbing et al., 2017). En ce moment, la quantité de données numériques doublerait tous les ans, une croissance exponentielle qui est difficile à apprécier intuitivement, mais qui, déjà, sous-tend et est sous-tendue par des entreprises parmi les plus puissantes et profitables au monde (« Fortune 500 Companies 2017 »,

s. d.), à savoir, les géants du secteur technologique ou « GAFAM » (Google, Apple, Facebook, Amazon, Microsoft). Souvent associée au discours sur le Big Data, on retrouve la notion d'intelligence artificielle (IA ou AI en anglais) ou de *machine learning* (apprentissage automatique, ou littéralement, apprentissage par une machine)¹. Le succès apparent du secteur technologique renforce la croyance que la solution se trouverait dans le *data*.

L'apparition de nouvelles technologies suscite à la fois la fascination et l'anxiété du public (Figure 3, (« Google Trends », s. d.). Avec des acteurs influents comme le cabinet de conseil McKinsey recommandant un déploiement massif du Big Data dans le secteur de santé (Manyika et al., 2011), et la volonté politique d'investir ces nouvelles technologies (Martos, 2017; Présidence de la République, 2018a, 2018b; Villani, 2018), il est peut-être temps pour une interrogation critique de ce qu'est le Big Data, et de ses implications pour la médecine et la psychiatrie. Le Big Data signifie-il réellement un bouleversement – pour le meilleur ou pour le pire – de la connaissance et de la pratique médicale, comme semblent le suggérer l'engouement et l'agitation médiatiques?

Pour répondre à cette question, cette thèse tentera d'abord de préciser ce que l'on entend par Big Data: sans arriver à une définition finale, il est possible d'esquisser les limites du concept en examinant le contexte dans lequel il est apparu, sa portée, ses caractéristiques et les outils qui y sont employés. Nous explorerons ensuite les utilisations actuelles du Big Data dans le domaine de la psychiatrie, à travers une revue sélective de la littérature. Enfin, ce contexte et ces exemples permettront d'éclairer les questions éthiques et épistémologiques que soulève l'utilisation du Big Data en médecine.

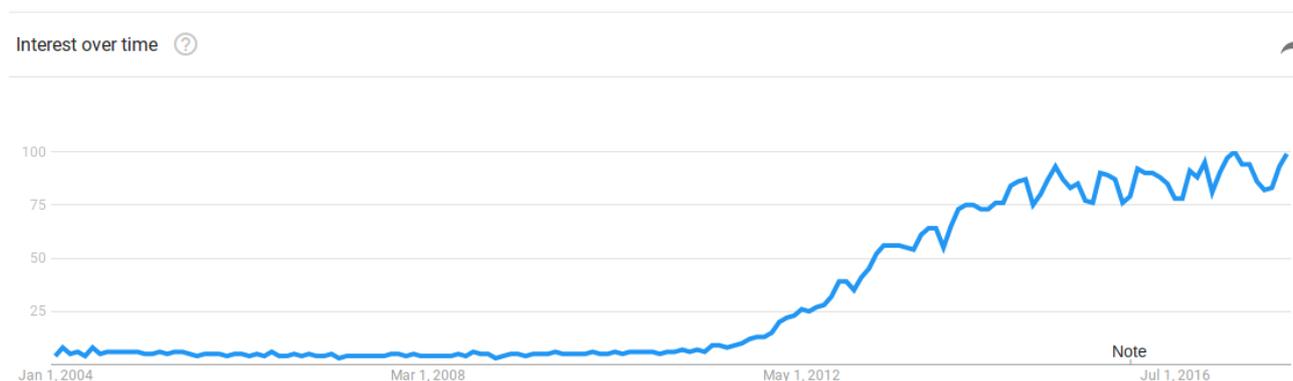


Figure 3 : Tendances des recherches sur Google pour « big data »

Courbe *Google Trends* de l'évolution des demandes pour « big data » sur son moteur de recherche, générée fin 2017 (« Google Trends », s. d.)

1 Une définition historiquement détaillée de l'intelligence artificielle ne sera pas abordée dans cette thèse (pour une discussion de cette définition, voir (Deutsch, 2012; Ganascia, 2010). Il suffit ici de clarifier que l'apprentissage automatique est une des formes d'intelligence artificielle.

I – Caractérisation du Big Data et des sciences des données

1. Cadre historique et propriétés principales

a. Pourquoi parle-t-on de Big Data ?

Ce qui distingue la nature des données numériques aujourd'hui et qui justifie l'appellation de Big Data est, bien entendu, en premier lieu leur quantité, leur volume sans parallèle dans l'histoire de l'humanité. S'y ajoutent la vitesse à laquelle elles sont générées et transmises, en temps-réel pour certaines ; ainsi que la variété de type et de format de données (structurées, non-structurées, texte, son, image, vidéo, localisation géographique, transactions financières, metadata etc). On parle de : « velocity, volume and variety » (Martin Hilbert, 2016). La différence, cependant, n'est pas seulement une différence d'échelle, mais aussi de nature. Ces données tendent à être (Lyon, 2014) :

- exhaustives, visant à capter des populations complètes et pas uniquement des échantillons comme en statistiques classiques ;
- détaillées (*fine-grained*) et liées à un identifiant individuel dont elles peuvent donner une représentation en « haute-définition » ;
- interconnectées de manière complexe, dans des bases qui peuvent être structurées ou non ;
- flexibles car facilement extensibles (*extensionality*) et évolutives (*scalability*), par la capacité d'élargir et de diversifier les bases de données rapidement.

Concrètement, cela signifie des bases avec plus d'entrées (exhaustives), mais surtout, plus de variables *individuelles* caractérisant chaque entrée (*fine-grained*), rapidement mises à jour, réactives et interconnectées. Pour emprunter au vocabulaire photographique, on aurait en même temps une vision panoramique et microscopique, instantanée et dynamique. Les bases deviennent d'une telle complexité qu'elles ne sont pas accessibles à un traitement statistique ou informatique classique, et nécessitent de nouveaux outils qui seront détaillés plus loin (1.3.). Certains auteurs argumentent que ce sont ces critères, plus que le critère de quantité absolue, qui distinguent le Big Data, et de ce fait ne qualifieraient pas les données très structurées d'un recensement de population

par exemple, aussi massives soient-elles, de Big Data. Mais cette position est discutable, car certains outils et traitements peuvent révéler des niveaux de complexité dans des données en l'apparence simples ou structurées.

De fait, la particularité des données massives tient aussi à la manière dont elles sont générées et utilisées. Il est utile d'insister ici sur le fait que ces données sont *interrogeables* et *interconnectées*, décuplant leur potentiel informatif. Par analogie, le contenu d'une vaste bibliothèque n'a de valeur que s'il y a un index permettant de chercher des livres selon certains critères et un bibliothécaire nous aidant à faire des liens inattendus et faire le tri. Alors que par le passé, l'enregistrement de données était fait sciemment (rédaction de manuscrits, tenue de registres d'état-civil, résultats d'une expérience ou d'une enquête etc), une partie non négligeable des données numériques sont générées de manière automatique, comme sous-produit d'activités électroniques, ou sans l'intention expresse du créateur de générer une base de données. Par exemple, un utilisateur de la plateforme de média social Twitter rédige des tweets pour communiquer ses opinions, discuter, créer un réseau social, mais ce-faisant contribue, sans en avoir consciemment l'intention, des données « volontaires » enregistrées et interrogeables, comprenant le contenu (ce qu'il a rédigé) et le contexte (l'heure et le lieu d'envoi, l'équipement utilisé). De manière semblable, les transactions bancaires et appels téléphoniques produisent des données automatiquement enregistrées sans intervention humaine, traces de l'utilisation de services numériques (Lyon, 2014).

En ce sens, le qualificatif de « Big » fait écho au *Big Brother* orwellien, son omniprésence, et les notions de contrôle (de la société et de la pensée). Mais il fait également penser à *Big Pharma* et *Big Ag*, l'idée de transformation d'un secteur économique, d'activité à une échelle industrielle, et de monopole. On voit donc que le sens de Big Data réside non seulement dans les caractéristiques propres aux données, mais aussi dans la manière particulière dont elles sont produites, peuvent être utilisées et sont liées entre elles, et dans leur place dans notre infrastructure économique et politique. Cette définition multiforme, faite d'un faisceau d'idées plutôt que de critères exclusifs, est à l'image de la complexité du vivant mentionnée en introduction. La figure 4 tente de schématiser le concept par un réseau de mots-clé.

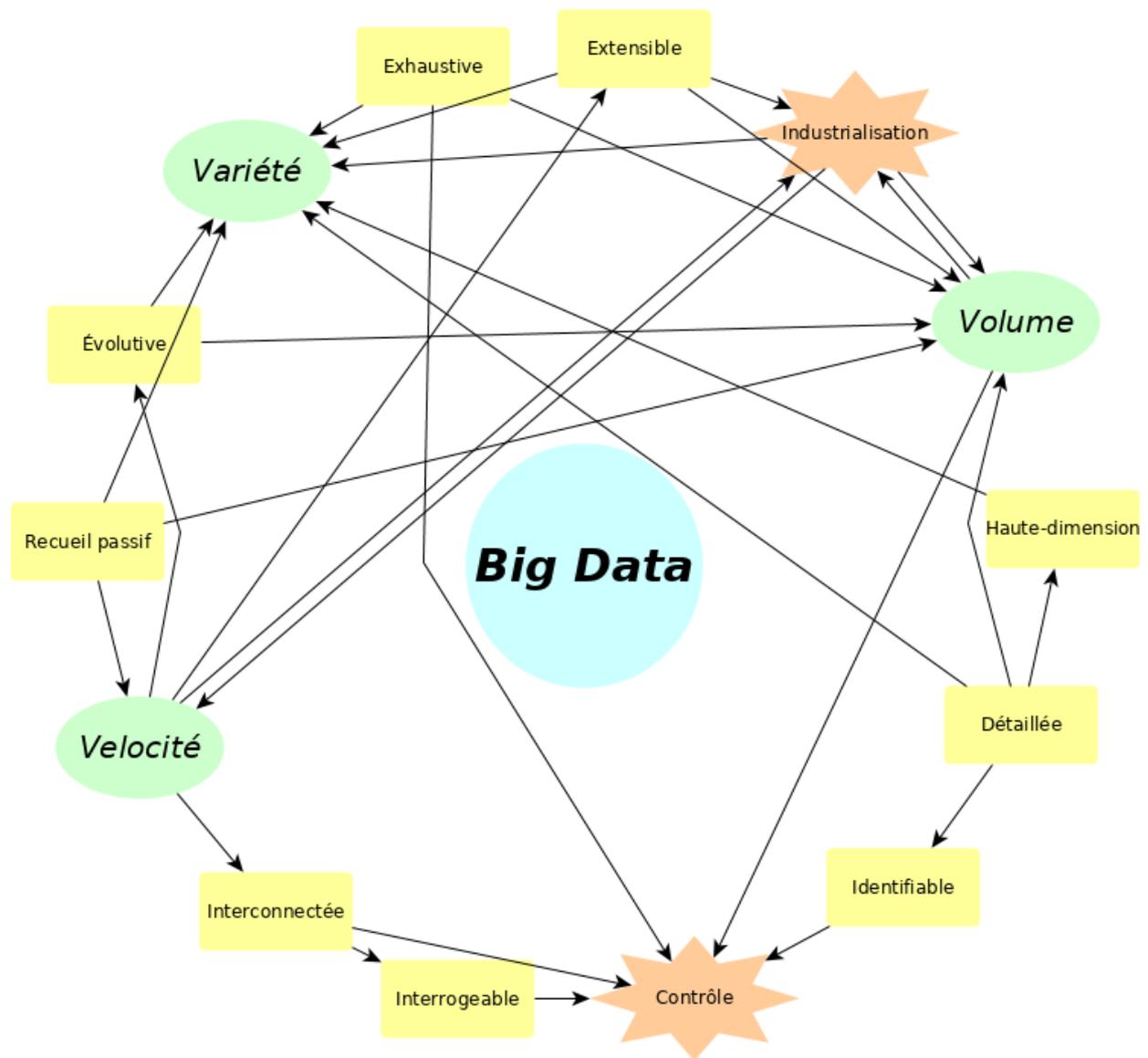


Figure 4 : Réseau de concepts définissant le Big Data

b. Comment en est-on arrivé là ?

Un bref aperçu historique de l'émergence des technologies numériques peut aider à mieux comprendre la topographie actuelle du Big Data et de la science des données, la place qu'occupe le Big Data dans l'imaginaire collectif et dans les faits, et les acteurs significatifs dans ce domaine. À travers les époques, les innovations en enregistrement et en transmission de données ont augmenté la quantité de données produites, la vitesse à laquelle elles sont générées et transmises, les capacités d'archivage et le type de

données enregistrables : que ce soit l'apparition de nouvelles méthodes d'encodage plus efficaces (chiffres romains vs chiffres arabes par exemple, système d'encodage binaire aujourd'hui), l'utilisation de supports plus compacts (papyrus vs tablettes d'écriture cunéiforme, DVD vs bande magnétique), l'automatisation de la reproduction de données (imprimerie vs manuscrits), le développement d'infrastructures et d'institutions dédiées à la transmission (poste, télégraphe, téléphonie) etc. On retrouve donc vitesse, volume et variété dans la trajectoire historique des technologies d'information. Ces développements diminuent en parallèle le coût associé à la production, au stockage et à la transmission de données (Roser & Ritchie, 2017).

L'essor du numérique correspond à une avancée remarquable dans cette trajectoire (M. Hilbert & Lopez, 2011). Historiquement, l'ordinateur est issu du concept de machine à calculer, outil mécanique qui simplifiait la réalisation d'opérations mathématiques, pour devenir une machine *programmée* à calculer, exécutant un programme, des instructions qui lui font réaliser des opérations mathématiques et logiques. Les premiers ordinateurs numériques étaient développés à l'époque de la Seconde Guerre Mondiale, principalement à usage militaire, avec un programme fixe réglé dans le câblage même de l'ordinateur, ou avec des programmes externes introduits par cartes perforées. Ils étaient suivis peu après par les premiers ordinateurs « modernes », dont l'architecture reposait sur des programmes enregistrés et programmés dans la mémoire de la machine (plutôt que dans la structure physique du câblage, ou sur des supports externes), les *stored-program computers* (Copeland, 2017; Haigh, 2011).

Ces premières itérations d'ordinateurs étaient mises au service d'applications pertinentes aux milieux où elles étaient nées – les secteurs militaire, industriel et public (gouvernement et universités) – et répondaient aux besoins de centralisation, d'efficacité et d'optimisation (Copeland, 2017; Ensmenger, 2012; Mahoney, 2005; Watson & Calhoun, 1960). La commercialisation des *mainframes*, ou ordinateurs centraux, dans les années 60 cimentait la place essentielle de l'informatique dans les grandes industries et entreprises, alliance qui figurait déjà depuis les prémices de cette discipline novatrice, comme l'atteste la création d'IBM, *International Business Machines*, au début du vingtième siècle (Ensmenger, 2012; Haigh, 2011). Avec l'augmentation exponentielle de la puissance de traitement des microprocesseurs (Figure 5, (Koomey, Berard, Sanchez, & Wong, 2011)), il devenait possible d'exploiter une masse croissante de données, permettant la réalisation de statistiques sur des échantillons de plus en plus larges. En parallèle aux capacités de traitement et de stockage, le développement d'interfaces utilisateur levait la nécessité

d'une expertise technique pour se servir de ces machines, augmentant ainsi le nombre de personnes pouvant générer et traiter des données.

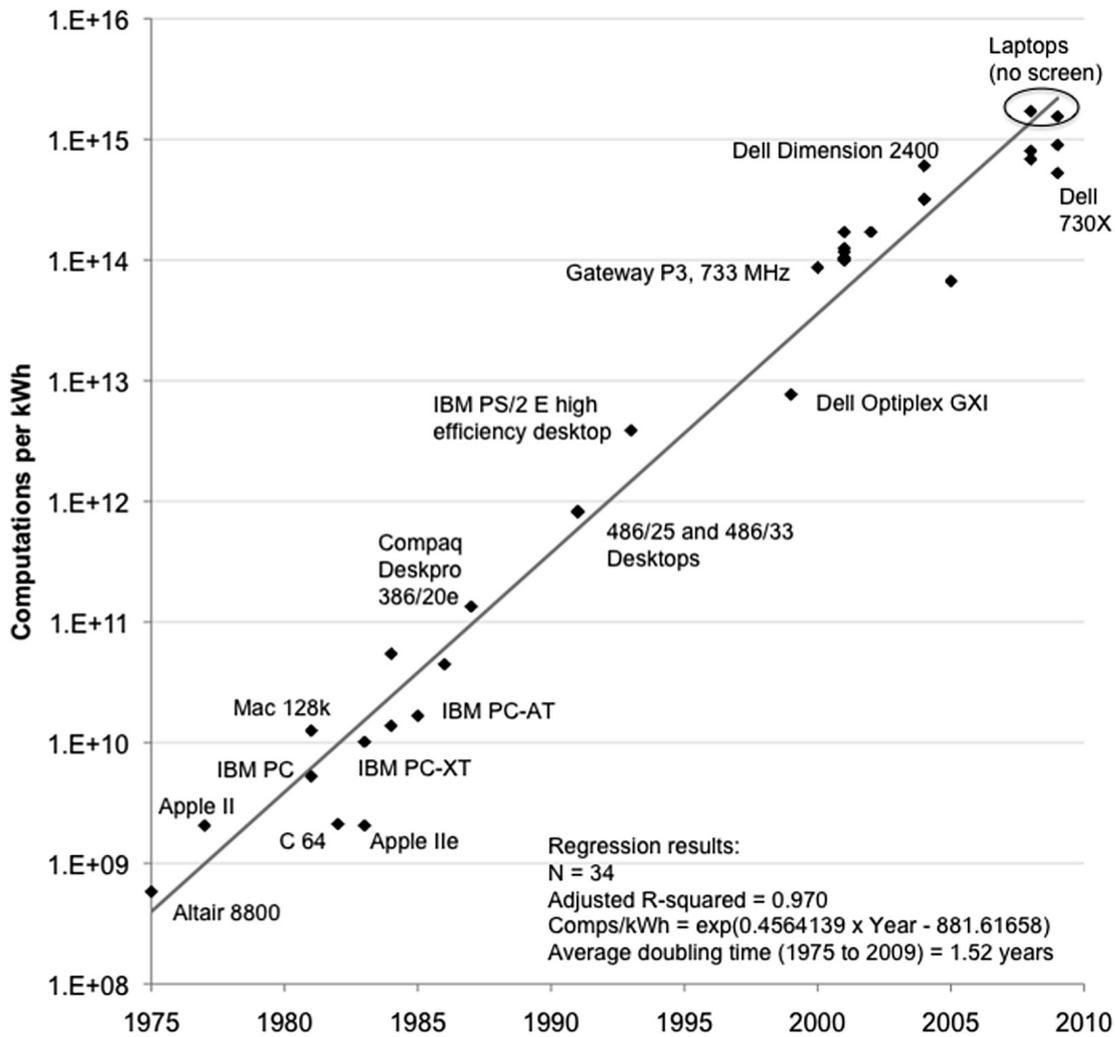


Figure 5 : Capacité de traitement des ordinateurs personnels de 1975 à 2010
(Koomey, Berard, Sanchez, & Wong, 2011)

Ces interfaces et la miniaturisation de l'équipement, notamment avec le circuit intégré, ont ouvert le chemin à la création de l'ordinateur personnel, dont le coût, la taille et la facilité d'utilisation le rendaient accessible au grand public (Haigh, 2011). Son apparition à partir des années 80, et la diffusion d'internet une décennie plus tard, ont démocratisé l'accès aux outils informatiques et créé un énorme réseau interconnecté de production et d'échange de données. Ces activités ne seraient donc plus l'apanage d'institutions publiques ou de groupes industriels, et par la même, tout deviendrait *data*, y compris les données en l'apparence anodines d'un profil sur un site de tchat ou sur une page web personnelle.

Lancé en 2007 par Apple, l'iPhone était révolutionnaire car il faisait du smartphone le conduit central de notre interaction avec le monde, en condensant en un seul appareil

compact, mobile, et facile d'utilisation, les fonctionnalités de téléphonie, messagerie, accès internet, photographie et musique (Markoff, 2007). Les ordinateurs portables avaient déjà permis une plus grande mobilité de l'informatique, et il existait des précurseurs qui combinaient certaines fonctionnalités comme les Palm Pilot puis le Blackberry, mais qui s'adressaient davantage à une clientèle professionnelle. L'iPhone, en revanche, était conçu pour tout public, avec un design attractif et ludique, et un écosystème d'applications à disposition. Il signalait le fait que le numérique était littéralement sur le bout des doigts, aussi indispensable qu'un téléphone au 20^{ème} siècle, intégré à nos activités professionnelles et personnelles, nos vies privées et publiques, et brouillant les frontières entre elles. Être connecté, consommer et produire des données devenait quelque chose que l'on pouvait faire *sans y penser*, sans devoir allumer un ordinateur, attendre qu'un modem se connecte etc.

Face à ce déluge de données, est né la *data science*. Bien qu'ayant des racines et une base mathématique communes, les disciplines des statistiques classiques et des sciences des données sont distinguées en pratique, les statistiques penchant davantage vers les mathématiques, et les sciences des données vers l'informatique (*computer science*) et la gestion (Mahoney, 2005). Il s'agit d'une division très grossière mais qui reflète l'évolution conjointe de l'informatique et du traitement des données, une évolution pragmatique entre des entreprises nécessitant des outils pratiques pour interpréter et visualiser leurs énormes bases de données et des ingénieurs inventant des solutions incorporées dans des logiciels informatiques (Breiman, 2001). Inversement, des innovations apparaissant dans des départements de sciences informatiques universitaires étaient rapidement absorbées ou transformées en entreprises. L'exemple phare est Google, dont le système « PageRank », inventé par Larry Page et Sergey Brin pendant leurs années de thèse, permettait d'exécuter un moteur de recherche largement plus performant que ceux qui le précédaient (Battelle, 2005). Leur projet universitaire initial était de comprendre l'architecture des liens entre pages Web, en s'inspirant du système de références et de citations dans les publications scientifiques. Ce sujet renvoyait à la question fondamentale de l'organisation de réseaux (théorie des graphes en mathématiques et informatique), mais de là sont nés une entreprise et un outil pratique, utilisé par des milliards de personnes, pour interpréter et trier l'information.

Ce contexte permet d'apprécier l'étendue de l'influence qu'exercent le secteur informatique et les géants du Web, sur le Big Data, ayant façonné ses outils, ses applications, son cadre conceptuel. Cette influence explique pourquoi, d'un point de vue

méthodologique, il est impossible d'aborder ce sujet par une recherche bibliographique classique, car il y a autant, voire davantage, d'informations cruciales dans des sources journalistiques économiques et technologiques. La communication et l'examen de nouvelles connaissances ou d'innovations techniques ne se fait pas forcément par des revues à comité de lecture, mais par d'autres moyens comme lors de congrès avec des démonstrations pratiques, ou au sein de l'archive de prépublication arXiv (« arXiv.org e-Print archive », s. d.). De même, ces travaux sont susceptibles d'être réalisés aussi bien dans des universités, des entreprises ou start-ups, qu'au cours de *hackathons* ou de concours de programmation en ligne, comme sur la plateforme Kaggle ouverte à tous (« Kaggle: Your Home for Data Science », s. d.). Cette différence de culture entre la biomédecine dont le gold standard est un modèle expérimental rigoureux, et le secteur numérique adepte du « bricolage » (*hacking, tinkering*), aura des implications épistémologiques qui seront abordées plus loin (III. 3.).

Le domaine de la santé n'a pas été délaissé par le secteur informatique. Parmi de nombreux acteurs privés utilisant le Big Data pour des applications médicales, on peut distinguer :

- Amazon : lancement en début d'année d'un projet conjoint avec la banque JPMorgan et la firme d'investissement Berkshire Hathaway pour transformer l'offre de soins pour leurs employés (Associated Press, 2018), entrée dans le réseau de distribution pharmaceutique aux États-Unis (Silverman, 2018) ;
- IBM : Watson Health développant une variété de programmes utilisables en médecine, santé publique, génomique et essais cliniques, comme Watson for Oncology utilisant l'intelligence artificielle pour l'aide à la décision clinique en cancérologie (Ross & Swetlitz, 2017) ;
 - Apple : Apple Health Records pour les dossiers médicaux électroniques, diverses fonctionnalités d'Apple Watch (détecteur de chute et de fibrillation atriale, (« Apple Watch Series 4 - Health », s. d.; Sheridan, 2018a)
- Verily Life Sciences, filiale du groupe Alphabet, maison mère de Google : développement d'appareils médicaux connectés pour surveillance et intervention, utilisation de Big Data en santé publique et en médecine de précision, partenaire du Precision Medicine Initiative (« Verily Life Sciences », s. d.) ;
- DeepMind, appartenant à Google : utilisation de l'intelligence artificielle dans plusieurs projets en partenariat avec des établissements de santé au Royaume-Uni et aux États-Unis (imagerie, dossiers cliniques : (« Researching for tomorrow », s. d.)), dont un

projet controversé en collaboration avec le Royal Free Hospital Trust à Londres (Naughton, 2017; Powles & Hodson, 2017) , et un outil diagnostique pour pathologies rétiniennes récemment publié (Fauw et al., 2018) ;

- Google Brain ou Research at Google : utilisation de *machine learning* en génomique et autres domaines de santé, travaux notables en analyse d'image automatique pour la rétinopathie diabétique (Gulshan et al., 2016) et la détection de cancers en histopathologie (Liu et al., 2017) ;
- Facebook et moteur de recherche Google : projets de dépistage et d'orientation pour la dépression et les idées suicidaires (Bach, 2017; « Building a Safer Community With New Suicide Prevention Tools | Facebook Newsroom », s. d.);
- de nombreuses start-up ou entreprises : WeDoctor en Chine (L. Y. Chen, 2018), Owkin (« OWKIN - Data Science and Artificial Intelligence for Medical Predictions », s. d.), Mindstrong Health (Dobbs, 2017), Qntfy (Qntfy Inc, s. d.), Medibio (« Medibio – Innovation In Mind », s. d.), Ginger.io (« Ginger.io », s. d.), PatientsLikeMe (« Live better, together! | PatientsLikeMe », s. d.), Woebot (« Woebot - Your charming robot friend who is ready to listen, 24/7 », s. d.) etc.

Cette liste est loin d'être exhaustive, mais laisse percevoir une volonté d'insertion à tous les niveaux de l'exercice clinique, de la recherche fondamentale aux interventions de santé publique, en passant par le dépistage, le diagnostic, le traitement et la surveillance.

Bien entendu, les institutions et autorités publiques ont également développé leurs politiques et projets de Big Data, à des fins économiques et sécuritaires, mais également sanitaires et sociales. En santé, les initiatives sont locales, nationales, régionales (Union Européenne, UE, par exemple) ou mondiales (Organisation Mondiale de la Santé, OMS), et concernent des bases de données déjà en existence, ou la création de nouvelles bases avec d'autres types de données. Un précurseur en France est le programme de médicalisation du système d'information (PMSI) et ses déclinaisons, offrant une vue sur les actes réalisés à l'hôpital, en plus des bases administratives d'actes remboursés par l'Assurance Maladie. Instauré en 2016, le Système National des Données de Santé (SNDS) regroupera à terme les bases PMSI, les bases de l'Assurance Maladie (base SNIIRAM), des données provenant des Maisons Départementales des Personnes Handicapées (MDPH), la base Inserm des causes médicales de décès (base CépiDC) et des échantillons de données provenant d'organismes d'Assurance Maladie complémentaire (« Qu'est-ce que le SNDS ? | SNDS », s. d.). Ces données, accessibles à des projets de recherche publics ou privés, présentent néanmoins davantage un intérêt

administratif que clinique. Par ailleurs, de nombreux jeux de données de santé publique sont mis à disposition par Santé publique France (ex-Institut de Veille Sanitaire) en partenariat avec des départements universitaires (InVS/Santé publique France, s. d.).

Sur un versant plus clinique, l'on peut trouver de nombreuses initiatives au sein du National Health Service (NHS) au Royaume-Uni, visant à créer des bases de données cliniques anonymisées ou pseudonymisées, à partir des dossiers médicaux informatisés. Certaines sont en partenariat avec des universités (« SAIL Databank - The Secure Anonymised Information Linkage Databank », s. d.) ; d'autres localisées au sein de groupements hospitaliers ou *trusts* (« Clinical Record Interactive Search (CRIS) », s. d.; Perera et al., 2016) d'autres encore, plus controversées, ont tenté de déployer un programme de partage systématique des données cliniques au niveau d'un territoire entier (care.data par NHS England, abandonné depuis, Boseley, 2016), ou se sont faites en partenariat avec une entreprise privée (DeepMind de Google et le Royal Free Hospital Trust, Naughton, 2017).

Au niveau Européen, le programme de recherche et d'innovation de la Commission Européenne, Horizon 2020, finance et encadre plusieurs projets Big Data en santé et en sciences, développant des outils informatiques en accès libre et extensibles pour aider la recherche en pharmacologie (« Health Pilot » Big Data Europe », s. d.; « Open PHACTS », s. d.), ou en établissant un dépôt standardisé de données pour les études financées par l'Union (« OpenAIRE », s. d.; « Zenodo - Research. Shared. », s. d.).

Les ressources développées et hébergées par la National Library of Medicine (NLM) et le National Center for Biotechnology Information (NCBI) sont un autre exemple de bases de données massives en accès libre mises à disposition par des organismes publics. La base de littérature biomédicale MEDLINE, les données issues de la génomique (bases de gènes, génomes, d'expression génique), ou de la biochimie (protéines, substances bioactives), sont toutes accessibles via le portail PubMed et NCBI (pubmeddev, s. d.). De même, la base Clinical Trials (« ClinicalTrials.gov », s. d.) regroupe des informations sur les essais cliniques, fournies par les investigateurs. En plus de ces bases déjà bien établies, le NIH investit dans des programmes supplémentaires de développement, de promotion, d'éducation et de recherches en Big Data, sciences des données et *machine learning*, à travers l'initiative Big Data to Knowledge (BD2K, (« About BD2K | Data Science at NIH », s. d.). Et l'un des volets de la Precision Medicine Initiative, intitulé « All of Us », est la constitution d'une base de données (démographiques, médicales, génétiques etc) regroupant au moins un million d'individus habitant aux États-

Unis (« National Institutes of Health (NIH) - All Of Us | National Institutes of Health (NIH) », s. d.).

Aux partenariats public-privé dont certains exemples sont cités plus hauts, s'ajoutent des associations, instituts et autres entités de la société civile qui tentent de promouvoir l'utilisation de Big Data. On peut citer notamment dans le domaine de la psychiatrie, l'organisation MQ Mental Health Research Charity basée au Royaume-Uni, qui finance et promeut une approche Big Data en recherche psychiatrique (« Mental Health Research Charity », s. d.).

c. Cadre légal

Un écosystème aussi peuplé peut s'organiser selon des principes très variés, en fonction de la distribution du pouvoir en son sein, des perspectives politiques, des priorités sociétales. Les ramifications du Big Data étant aussi profondes que larges, la volonté de régulation de ce secteur se comprend comme une volonté de réduction des risques, de capitalisation d'une ressource précieuse et de contrôle d'une force potentiellement envahissante. De fait, les cadres juridiques et l'étendue de régulation diffèrent fondamentalement entre pays et entre régions du monde, en fonction des spécificités de leurs cultures politiques. En la matière, le contraste entre l'opacité et la laxité réglementaire aux États-Unis, et les efforts de réglementation européens illustre bien le degré de divergence (Executive Office of the President, 2014; Executive Office of the President & President's Council of Advisors on Science and Technology, 2014). La législation sur les données peut recouvrir différents domaines, selon qu'il s'agisse de données ayant trait à la personne humaine ou non. Ainsi, certaines données peuvent relever de textes sur la propriété intellectuelle ou les droits d'auteur par exemple, alors que d'autres, les données personnelles, qui ont un intérêt particulier en santé, relèvent de textes spécifiques.

En France, des textes encadrant l'utilisation de données personnelles, définies comme « toute information identifiant directement ou indirectement une personne physique (nom, numéro d'immatriculation, numéro de téléphone, photographie, date de naissance, commune de résidence, empreinte digitale...) », existent déjà depuis plusieurs décennies (« Glossaire | CNIL », s. d.). Les données de santé à caractère personnel sont, de surcroît, considérées comme données sensibles. La Commission Nationale de

l'Informatique et des Libertés (CNIL), créée en 1978 par la loi Informatique et Libertés (Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, s. d.) est une autorité administrative indépendante chargée de réguler les traitements numériques des données personnelles. Elle a pour mission d'informer et d'accompagner les professionnels dans leur mise en conformité, et les particuliers dans l'exercice de leurs droits à maîtriser leurs données personnelles (« Les missions | CNIL », s. d.). Elle a l'autorité de contrôler et sanctionner les violations d'utilisations de données personnelles. Elle mène des campagnes de formation et de sensibilisation du public sur les questions de la sécurité numérique et de la protection de la vie privée. Elle essaie d'anticiper l'évolution des nouvelles technologies, et les réponses à y apporter, pouvant être consultée pour un avis ou en menant une réflexion publique sur un sujet, comme sur la question des algorithmes (« Débat public sur les algorithmes et l'intelligence artificielle », s. d.).

En matière de santé, la CNIL, en collaboration avec d'autres agences comme l'agence française de la santé numérique (ASIP Santé), élabore des référentiels et des procédures d'agrément pour les organismes collectant, hébergeant et échangeant des données personnelles de santé. Sur le versant recherche, la plupart des projets utilisant des données personnelles de santé font l'objet, au minimum, d'une déclaration à la CNIL, et nécessitent pour certains d'entre eux, une autorisation de la CNIL, en plus des autres autorisations nécessaires selon le cas (comité de protection des personnes CPP, pour les recherches interventionnelles humaines ; institut national des données de santé INDS, pour l'accès aux données du SNDS).

La loi Santé de 2016, qui définit le cadre réglementaire du SNDS, vise explicitement à « créer les conditions d'un accès ouvert aux données de santé » (LOI n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé - Article 193, 2016), modifiant certaines dispositions de la loi Informatique et Libertés, simplifiant l'accès et le partage de données dans le cadre des soins, clarifiant la place du consentement des personnes à l'utilisation et la réutilisation de leurs données, définissant l'ouverture des données de santé à des fins de recherche, et créant l'INDS comme « portail unique » pour la mise à disposition de ces données.

Au delà des données de santé, les droits fondamentaux garantis par la loi Informatique et Libertés sont (« Comprendre vos droits | CNIL », s. d.) :

- droit d'accès : possibilité de demander directement à un organisme s'il détient des données personnelles sur soi et d'accéder à l'intégralité de ces données ;
- droit de rectification : demander de corriger des informations inexacts sur soi ;

- droit d'opposition : possibilité de s'opposer à figurer dans un fichier ou que ses données personnelles soient diffusées, transmises ou conservées ;
- droit au déréférencement : saisir les moteurs de recherche pour déréférencer une page web associée à une personne, pour des raisons d'atteinte à la vie privée ;
- droit d'accès aux fichiers de police, de gendarmerie, de renseignement, d'administration fiscale : possibilité de demande d'accès aux fichiers concernant le demandeur, par l'intermédiaire de la CNIL. La demande est donc indirecte et le droit d'accès à ces fichiers est soumis à l'accord de l'autorité administrative en question.

En contrepartie, les organismes traitant des données personnelles ont certaines obligations, permettant d'assurer un usage légitime et respectueux de la vie privée (« Les principes clés de la protection des données personnelles | CNIL », s. d.). Ils doivent :

- définir les finalités, c'est-à-dire les objectifs du fichier, à quoi vont servir les données, limitant ainsi la manière dont les données peuvent être utilisées ou réutilisées ;
- minimiser la collecte en vérifiant la pertinence des données collectées, ne capturant que celles qui sont strictement nécessaires à la réalisation de l'objectif ;
- limiter la durée de conservation des données qui doivent être supprimées une fois que la finalité est atteinte, en tenant compte des éventuelles obligations à conserver certaines données ;
- respecter les droits des personnes en les informant obligatoirement de la collecte de données dont ils font l'objet, de la nature des données collectées, leur finalité et la durée de conservation, ainsi qu'en mettant en place des procédures facilitant l'exercice des droits d'accès, de rectification et d'opposition ;
- garantir la sécurité et la confidentialité des données en utilisant toutes les mesures nécessaires, de manière adéquate au niveau de risque (données sensibles, type d'utilisation et de finalité etc).

L'adoption au niveau européen du Règlement Général sur la Protection des Données (RGPD), applicable depuis Mai 2018, harmonisera la législation au niveau de l'UE (« Key Changes with the General Data Protection Regulation », s. d.; « Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 | CNIL », s. d.). Il reprend des principes déjà ancrés dans la loi Informatique et Libertés, en renforce certains et en introduit d'autres, aboutissant à un cadre juridique qui priorise la protection de la vie privée, la maîtrise des particuliers sur leurs propres données et la responsabilisation des acteurs traitant des données. Une attention particulière y est portée pour les données

personnelles de santé. Les points forts et nouveautés du RGPD concernent (Figure 6, (« Plus de droits pour vos données ! | CNIL », s. d.) :

- les droits des particuliers : portabilité des données (possibilité de récupérer, de réutiliser, de transférer leurs données personnelles pour répondre à leurs propres besoins), meilleure lisibilité sur l'utilisation et la collecte des données, simplification de l'exercice des autres droits, meilleure protection des mineurs, consécration du droit à l'oubli (déréférencement) ;
- les obligations des traitants : renforcement des sanctions en cas de violations, obligation d'une information claire et aisément accessible aux personnes concernées par un traitement de données ;
- la portée et le cadre : extension du champ d'application à toute personne résidant dans l'UE, coopération renforcée entre les autorités de protection des données, principe de protection des données dès la conception d'un produit/service et par défaut (*privacy by design, privacy by default*), meilleure reconnaissance du risque lié au traitement de données (études d'impact sur la vie privée, obligation de notification en cas de violations de données personnelles, droit à réparation d'un dommage matériel ou moral).

La législation aborde aussi la question de la décision individuelle automatisée, c'est-à-dire les décisions issues d'un traitement automatisé, par un algorithme, sans intervention humaine, comme le profilage. Ces décisions peuvent avoir des conséquences significatives, voire même juridiques, pour la personne concernée. Le texte du RGPD insiste sur le besoin de sauvegarder les droits et les libertés des personnes en permettant la contestation d'une décision automatisée, et en limitant ces décisions dans le cas de données sensibles comme les données de santé.

Ces différents cadres législatifs ont le mérite d'être basés sur un principe fort de protection des libertés individuelles, de la vie privée, et donc de l'autonomie des personnes. Les données de santé sont systématiquement mentionnées avec une attention particulière et une reconnaissance de leur caractère hautement sensible. Les textes sont rédigés avec un souci de précision mais aussi d'équilibre, laissant des marges de manœuvre. Cependant, la mise en application de ces lois, notamment du RGPD, s'annonce déjà très compliquée, si ce n'est parce qu'elle implique potentiellement un vaste nombre d'acteurs numériques au niveau international. Cela reste à voir si les idéaux incarnés par la législation française et européenne seront réalisés en pratique et

deviendront aussi une référence internationale, ou si au contraire, ils feront l'objet d'un « bricolage » ouvrant des brèches au niveau de leurs points faibles.



Figure 6 : Infographie résumant les points forts du RGPD (« Plus de droits pour vos données ! | CNIL », s. d.)

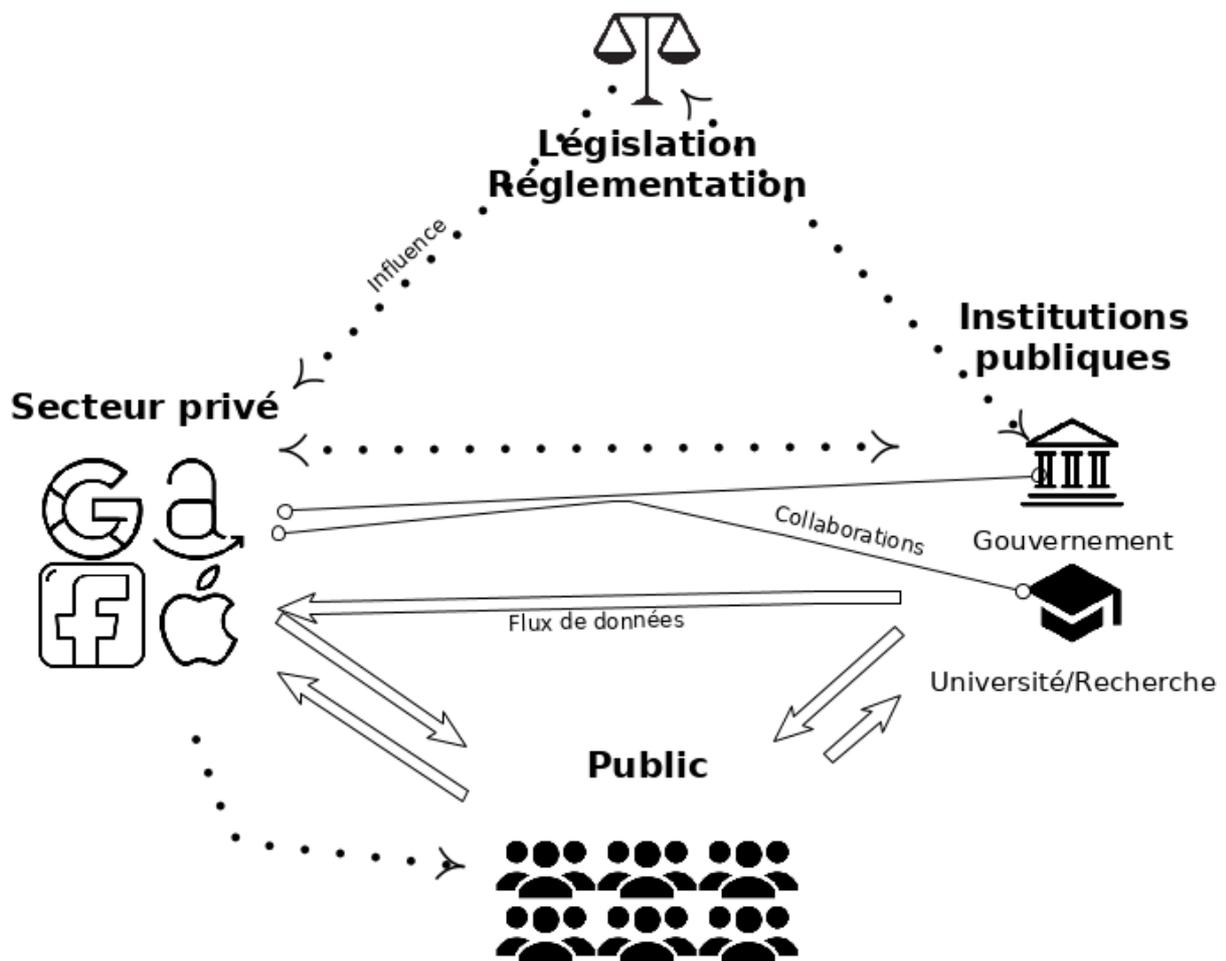


Figure 7 : Écosystème du Big Data

Acteurs majeurs, directions du flux de données, de l'influence et des collaborations (graphique réalisé avec l'aide d'icônes sous licence Creative Commons de (Boyko, s. d.; youtube.com/AlfredoCreates, s. d.)

2. Spécificités des données en santé et en recherche biomédicale

Si les données de santé sont aussi sensibles, elles présentent également d'autres défis à leur utilisation. Ayant trait à l'intimité des personnes, à leur vulnérabilité, elles nécessitent une protection particulière, qui est représentée par le secret médical dans la relation soignant-soigné, et qui doit trouver son équivalent dans le traitement numérique. La législation insiste sur la nécessité de sécuriser les données, mais nous verrons plus loin (III) des exemples d'échecs retentissants en termes de protection des données.

Bien que les bases de santé publique et les bases administratives de santé soient souvent centralisées (SNDS, CépiDC à l'Inserm pour les causes médicales de décès, diverses bases de Santé publique France, base publique de médicaments sur data.gouv.fr), la plupart des données personnelles de santé, y compris pour une même personne, sont éparpillées dans les systèmes informatiques des divers établissements et praticiens de santé, sans compter les données non numérisées qui gisent dans les dossiers médicaux « papier ». Cette fragmentation complique le traitement en masse des données, mais en cela, elle est aussi une protection contre la captation totale d'informations sensibles et la constitution de « profils » numériques.

C'est précisément contre cette protection que les géants du Web mettent en place des stratégies, comme l'acquisition de start-up et de sites populaires (achat d'Instagram par Facebook, ou de YouTube par Google) ou l'obligation d'un enregistrement nominatif pour accéder à différents services (obligation d'entrer un compte Google ou Apple pour configurer un téléphone, commentaires sur certains sites réservés uniquement aux utilisateurs de Facebook). En effet, une des clés du traitement de masse en Big Data est l'existence de données interconnectées par un même identifiant (*individual-level linkage*), postulat à la base de la Precision Medicine Initiative et des autres initiatives Big Data gouvernementales comme le SNDS, SAIL Databank (Secure Anonymised Information Linkage Databank) etc. Sur ce point, les données médicales représentent un défi et une opportunité, et motivent la mise en place de projets massifs de collecte comme la base de génomes « All of Us », et ses équivalents ailleurs dans le monde (Auffray et al., 2016). Mais même avec des bases massives connectées et identifiées, les données ayant trait à la santé présentent certaines particularités dont les principales seront détaillées ci-dessous.

a. Données génétiques et d'imagerie : haute précision et haute dimensionalité

Les données d'analyse du génome (ADN) ou du transcriptome (ARN), ou d'imagerie cérébrale notamment en imagerie par résonance magnétique (IRM), sont utilisées en clinique, mais surtout fréquemment en recherche biomédicale en psychiatrie et neurosciences. Elles sont ainsi au carrefour entre recherche fondamentale et application clinique. En oncologie, la traduction de données génétiques en décisions cliniques se fait

déjà de manière courante (Auffray et al., 2016), et nombreux sont ceux qui espèrent pouvoir faire de même en psychiatrie.

Ces données sont l'exemple typique de données relativement structurées, mais à très haute-dimension (*high-dimensionality*), contenant des centaines de milliers, voire des millions, de variables (*features* : ici les paires de bases, les SNPs, les voxels...) pour un échantillon. La haute-dimensionnalité correspond à toute situation où le nombre de variables excède le nombre d'échantillons, ce qui est valable pour une étude de cinq échantillons/sujets avec une analyse de dix variables, et l'est d'autant plus pour les données d'imagerie et de génétique. Dans ces cas, les outils statistiques classiques, y compris les analyses multivariées, atteignent leurs limites et certains ne peuvent pas être utilisés pour l'interprétation des données (James, Witten, Hastie, & Tibshirani, 2013). Lorsqu'on sélectionne a priori les variables à analyser (par exemple en choisissant de comparer des régions d'intérêt en imagerie, ou des gènes candidats), on contourne ce problème car on réduit les dimensions (Breiman, 2001). On risque alors une perte d'information, et cela nécessite de toute manière, une hypothèse, un modèle de départ. Mais si on souhaite comparer les échantillons sur la totalité des variables, présumant par exemple, qu'il existe des corrélations jusque-là négligées et de nouveaux modèles à découvrir dans ces données, on est obligé de gérer la haute-dimensionnalité. On peut essayer d'augmenter le nombre d'échantillons jusqu'à excéder largement le nombre de variables. La diminution du coût d'analyse génomique et la simplicité du prélèvement facilitent en effet la collecte d'un nombre croissant d'échantillons génétiques. Cependant, une telle augmentation d'échelle n'est actuellement pas possible en imagerie, et on peut sérieusement se demander si elle est souhaitable en génétique. Ces difficultés amèneront finalement au recours à des outils d'analyse spécifiques aux situations à haute-dimension (voir chapitre I.3.b, Breiman, 2001; James et al., 2013).

Même avec des outils plus adaptés, lorsqu'on manipule un si grand nombre de variables, d'autant plus dans un système biologique dégénèrescent et redondant, on se heurte au problème de la corrélation de nombreuses variables entre elles, aussi appelé « multi-colinéarité » (James et al., 2013). Ceci implique qu'il est impossible de savoir, à partir des seules données, aussi vastes soient-elles, exactement quelles sont les variables directement associées et réellement prédictives, d'un point de vue physiopathologique, du résultat recherché. On pourra ainsi obtenir plusieurs combinaisons de variables prédictives du statut malade ou non-malade, sans que l'une soit exclusivement « vraie ». Pour prendre un exemple fictif et caricatural, on pourrait avoir une analyse génomique d'une

population urbaine identifiant dix allèles prédictifs de schizophrénie, qui correspondent à des gènes impliqués dans la neurogenèse et d'autres associés au complexe majeur d'histocompatibilité commun parmi certains groupes ethniques migrants. Cette combinaison de dix allèles se révèle utile pour prédire qui sera atteint de schizophrénie, mais nous dit-elle que le système immunitaire est impliqué dans le mécanisme pathologique, que le groupe ethnique ou le statut de migrant est lié à la maladie, ou que ce sont les anomalies du système nerveux qui sont prédictives, les autres associations étant le fruit du hasard ? Ou encore, on pourrait avoir une analyse génomique qui identifie dix allèles prédictifs d'une meilleure réponse à un antidépresseur, correspondant à des gènes du métabolisme hépatique et des gènes impliqués dans la transmission GABA-ergique et le rythme circadien. Cette combinaison de dix allèles se révèle utile pour prédire quel patient répondra mieux à un certain antidépresseur, mais nous dit-elle que c'est la variabilité pharmacocinétique ou phénotypique qui détermine la réponse ? C'est en quelque sorte, le pendant du concept de facteur confondant, à un degré supérieur compte-tenu du nombre de variables en jeu.

Malgré la haute-dimension, les données de génétique et d'imagerie ont généralement une faible résolution temporelle, offrant un cliché statique à un temps donné, ou à intervalles fixes si les prélèvements sont répétés. L'IRM fonctionnelle, l'analyse du connectome en imagerie, du transcriptome ou du protéome en biologie, ajoutent une dimension temporelle en permettant de visualiser des processus dynamiques limités, des changements chronologiquement associés à un stimulus ou une perturbation. Pour autant, ce type de données, surtout l'imagerie, ne servira vraisemblablement pas à un suivi continu.

b. Dossiers médicaux informatisés et publications : le défi du texte libre et de la fragmentation

Les données des dossiers médicaux ont une plus grande richesse temporelle que les données précédentes, avec la possibilité de suivis longitudinaux, des intervalles d'observation variés et des données historiques sous la forme d'antécédents. Leur enregistrement faisant partie de la pratique clinique, ces données sont déjà largement disponibles et faciles à accumuler, notamment lorsque des systèmes de dossiers médicaux informatisés sont en place. Le degré de précision des observations cliniques offre un

niveau élevé de détail, et la variété des données que l'on trouve dans un dossier médical (informations administratives, démographiques, observations médicales et paramédicales y compris antécédents familiaux concernant d'autres individus, examens complémentaires etc) en font des jeux de données à haute-dimension, comme ceux décrits ci-dessus.

À la haute-dimension et la multi-colinéarité, s'ajoutent d'autres défis de traitement et d'interprétation. La fragmentation des données au sein de plusieurs systèmes, et sous plusieurs formats de fichiers pas toujours compatibles entre eux, freine l'agrégation des données et leur organisation en une base facilement interrogeable, à la fois pour trouver des informations sur les patients individuels, et pour faire des analyses sur un grand nombre de personnes. En pratique, au sein d'un même établissement, il existe encore des dépôts différents pour les observations cliniques, les résultats de biologie, l'imagerie et les prescriptions, empêchant une vue d'ensemble de la situation médicale d'une personne. Ce problème est multiplié lorsque plusieurs praticiens ou établissements sont en jeu.

De plus, malgré l'existence de documents structurés comme les compte-rendus d'hospitalisation et les codages PMSI ou CIM (Classification Internationale des Maladies, (Organisation mondiale de la santé, 2009), la richesse et la granularité gisent sous des formes non-structurées, dans les observations des soignants en texte libre (Luo, Wu, Gopukumar, & Zhao, 2016; Meystre, Savova, Kipper-Schuler, & Hurdle, 2008). Dans ce type de texte, le format, mais aussi la syntaxe, peuvent être non-structurés, à travers l'utilisation d'un style télégraphique, l'absence de ponctuation, l'emploi d'abréviations multiples et de termes particuliers qui, malgré un lexique médical commun, peuvent être source d'ambiguïté. L'exploitation de ces données nécessite donc non seulement un système capable d'interpréter des sources en langage naturel/ordinaire, mais de gérer les particularités de ce langage tel qu'il est utilisé dans le contexte clinique.

La littérature médicale et scientifique constitue une base de données moins sensible car ne traitant pas de données personnelles. Elle est déjà facilement interrogeable pour la recherche de publications individuelles, grâce à des moteurs de recherche et des bases en ligne étendues. Mais l'exploitation de cette vaste source à des fins d'agrégation, d'extraction et de comparaison de données, se fait encore manuellement, comme en atteste la méthodologie des travaux de revue, de méta-analyse, de conférence de consensus et d'élaboration de recommandations (sélection des publications, lecture d'abstracts puis de papiers, compilation des données et demande de données supplémentaires auprès des auteurs). De fait, l'automatisation de ces processus et l'exploitation à plus grande échelle posent des défis semblables à ceux rencontrés avec

les dossiers médicaux. Bien qu'elles aient un format stéréotypé plus structuré que les dossiers médicaux, les informations que contiennent les publications se trouvent aussi au sein d'un texte en langage naturel, et certaines sont, de surcroît, bloquées derrière un *paywall* réservant le contenu aux abonnés. La limitation d'accès, la fragmentation des données, la nécessité d'extraire des données non-structurées à partir du langage ordinaire et de les organiser de manière intelligible, sont parmi les caractéristiques principales et les plus épineuses du traitement en masse de données de santé, et appellent l'invention de nouveaux outils et de nouveaux cadres de traitement.

c. Données collectées par smartphone ou appareils connectés

L'ubiquité des smartphones et leur simplicité – relative – d'utilisation, en font un outil attrayant pour collecter des informations individuelles et mettre en place des interventions en temps réel. Théoriquement, on aurait une résolution temporelle presque illimitée, avec des données collectées et transmises en continu et en temps réel, par un appareil toujours à portée de main. Le smartphone fournirait des données plus directes, et prétendument plus objectives, ne dépendant pas de la subjectivité d'un interrogatoire et d'une relation interpersonnelle avec un soignant. Centralisées au sein d'un appareil personnel, et facilement transmissibles d'un appareil à un autre, ces données ne nécessiteraient pas des outils complexes d'extraction et d'agrégation, contrairement aux dossiers médicaux par exemple.

La plupart des smartphones ont des fonctionnalités intégrées qui enregistrent déjà régulièrement ou peuvent enregistrer, avec l'aide d'applications spécifiques, les données suivantes (Onnela & Rauch, 2016) :

- géolocalisation par GPS, par triangulation d'antennes-relais, ou par adresse Wifi,
- volume de données échangées au cours du temps,
- consommation de batterie au cours du temps, avec le détail par application,
- historique de sites internet visités,
- temps d'émission ou de réception, destinataire ou émetteur et durée d'un message ou d'un appel (métadonnées),
- contenu d'un message, enregistrement audio d'un appel,
- enregistrement audio en continu de tous les sons dans l'environnement proche du téléphone,

- données d'accéléromètre détectant l'accélération et l'orientation du téléphone (possibilité de détecter si le téléphone est en mouvement),
- mémoire cache et sauvegarde de données utilisées ou échangées par les applications installées.

De nombreux appareils connectés déjà commercialisés ou en cours de développement, peuvent pratiquer des mesures supplémentaires qu'ils transmettent ensuite à un smartphone par une connexion sans-fil Bluetooth ou WiFi pour archivage, interprétation ou échange. Le Fitbit, par exemple, commercialisé comme aide à la pratique sportive, pourrait selon ses fabricants, mesurer la fréquence cardiaque, enregistrer le niveau d'activité physique et interpréter les différentes phases de sommeil (« Technology - Fitbit », s. d.), bien que la précision des mesures cardiaques soit remise en question (Wang et al., 2017). Récemment, la FDA (*Food and Drug Administration*, équivalent américain de l'ANSM, agence nationale de sécurité du médicament) a approuvé la commercialisation d'une pilule d'Abilify (aripiprazole) avec détecteur de prise intégré, qui émet un signal électrique une fois que la pilule se trouve dans l'estomac. Ce signal est détecté par un patch collé sur la paroi thoraco-abdominale, qui envoie ensuite l'heure et la date d'ingestion à une application correspondante installée sur le smartphone du patient (Belluck, 2017). Ces appareils et détecteurs connectés miniaturisés, formant ce que l'on appelle l'Internet des Objets (*Internet of Things*, IdO ou IoT), étendent les fonctionnalités d'un smartphone ou d'un ordinateur, avec des applications indirectement ou explicitement médicales, augmentant le nombre et le type de données enregistrables.

Un autre exemple très récent est l'Apple Watch 4 dévoilée en septembre 2018, avec un gyroscope et un accéléromètre pour détecter des chutes, et des électrodes pour enregistrer un électrocardiogramme (ECG) afin de détecter la survenue de fibrillation atriale (« Apple Watch Series 4 - Health », s. d.; Center for Devices and Radiological Health, s. d.; Sheridan, 2018a). À la différence des appareils précédents qui doivent être portés sur soi, d'autres ont pour but de détecter des phénomènes à distance. Vital-Radio, en développement au Massachusetts Institute of Technology, se base sur les interférences sur un signal WiFi, dus à des obstacles ou des mouvements dans l'entourage du signal (Adib, Mao, Kabelac, Katabi, & Miller, 2015). L'appareil génère un faible signal WiFi, et détecte ensuite des signaux en retour (comme un radar ou une échographie), qu'il transmet à un ordinateur pour traitement. L'algorithme de traitement du signal serait alors capable d'extraire les interférences dues aux mouvements respiratoires et aux vibrations thoraciques des battements cardiaques, pour obtenir un monitoring continu, à distance,

des fréquences cardiaques et respiratoires. L'appareil fonctionnerait également avec plusieurs personnes présentes (deux à trois) et dans une pièce voisine, et des travaux récents montreraient une capacité à détecter les mouvements dans une ou plusieurs pièces (*NIH VideoCast - Artificial Intelligence: Collecting Data to Maximize Potential*, 2018).

Seuls les appareils et logiciels relevant du statut de dispositif médical sont soumis à une réglementation spéciale avant leur commercialisation, et une surveillance par l'ANSM. Selon le code de la santé publique : « On entend par dispositif médical tout instrument, appareil, équipement, matière, produit [...], ou autre article utilisé seul ou en association, y compris les accessoires et logiciels nécessaires au bon fonctionnement de celui-ci, destiné par le fabricant à être utilisé chez l'homme à des fins médicales [...]. Constitue également un dispositif médical le logiciel destiné par le fabricant à être utilisé spécifiquement à des fins diagnostiques ou thérapeutiques. » (Code de la santé publique - Article L5211-1, s. d.)

« Ces dispositifs sont destinés à être utilisés à des fins :

1° De diagnostic, de prévention, de contrôle, de traitement ou d'atténuation d'une maladie ;

2° De diagnostic, de contrôle, de traitement, d'atténuation ou de compensation d'une blessure ou d'un handicap ;

3° D'étude, de remplacement ou de modification de l'anatomie ou d'un processus physiologique ;

4° De maîtrise de la conception. » (Code de la santé publique - Article R5211-1, s. d.)

Malgré ces définitions, des appareils ou logiciels commercialisés ostensiblement pour des motifs non-médicaux, mais dont l'utilisation serait détournée à des fins médicales, se retrouveraient dans une sorte de vide réglementaire, un décalage entre l'usage réel et l'intention déclarée. L'absence de référentiels existants pour certains types de données, comme l'accélérométrie, limite encore l'interprétation rigoureuse de ces données et la formulation de conclusions significatives et généralisables. La surveillance continue de personnes généralement en bonne santé, soulève des questions déontologiques et éthiques, et générera pour le moins une vaste quantité de données que l'on ne sait pas encore interpréter, avec un risque élevé de fausses alertes. En effet, sur la base des données fournies à la FDA pour l'autorisation de commercialisation, le détecteur de fibrillation atriale d'Apple Watch aurait une valeur prédictive positive d'environ 45 %, donc une proportion non négligeable de faux positifs (Sheridan, 2018a). Enfin, la

vulnérabilité de données personnelles constamment transmises via de multiples connexions à longue et courte-distance met en relief la question de la sécurité des données, et la nécessité de systèmes de protection et de chiffrement pratiques à utiliser.

d. Données de santé publique

Contrairement aux données décrites ci-dessus, les données de santé publique sont souvent déjà organisées en grandes bases structurées, exploitables, interrogeables. L'adoption du numérique a bien entendu facilité la compilation de ces bases, en réduisant le coût et augmentant la capacité d'archivage, en accélérant la vitesse de collecte et d'entrée de l'information. Mais au delà de ces avancées notables, une approche Big Data amène d'autres manières d'exploiter, d'interpréter et de visualiser les données, éclairant les questions de santé publique sous des angles différents :

- accès aux données facilité par des plateformes rassemblant plusieurs jeux de données en un lieu repérable, et des portails proposant plusieurs options d'accès (le SNDS, ou data.gouv.fr par exemple),
- multiplication et simplification des modalités d'entrée de données (formulaires en ligne ou par application smartphone, agrégation automatique de données administratives), augmentant potentiellement le nombre d'acteurs pouvant contribuer des données, et la réactivité du recueil de données (ajout et suppression d'indicateurs recueillis en fonction des besoins, collecte en temps réel ou à intervalles de quelques jours),
- combinaison et comparaison de bases de santé multiples, ou des bases de santé avec d'autres bases de statistiques publiques (socio-économiques, par exemple),
- outils de cartographie simples d'utilisation, pour une visualisation géographique des données,
- plus large dissémination des résultats par des supports multiples (texte, vidéos, graphiques interactifs, directement en ligne ou par la presse etc), pouvant s'adapter aux différents niveaux d'expertise du public visé.

Le réseau Sentinelles de surveillance continue et de modélisation de certaines maladies infectieuses et de suicide, coordonné par une unité de recherche Inserm de l'Université Paris 6 et Santé Publique France, est un bon exemple de la mise en application de ces fonctionnalités ((« Réseau Sentinelles > France > Le réseau Sentinelles », s. d.), Figure 8). Un réseau de plus d'un millier de médecins généralistes et

de pédiatres libéraux, répartis sur tout le territoire, transmettent toutes les semaines, par une plateforme sécurisée en ligne, des données sur les patients vus en consultation, à propos de plusieurs indicateurs de santé (neuf indicateurs actuellement, mais la flexibilité du système permet de les modifier si besoin). Ces données sont validées et traitées, permettant de formuler des estimations nationales et des modèles du comportement de ces maladies au cours du temps. Les résultats sont disponibles sous la forme de tableaux de données en chiffres, présentant un intérêt pour des chercheurs souhaitant les réutiliser ou les combiner avec d'autres jeux de données, ou sous la forme de cartes et de graphiques accessibles à tous. Le réseau s'étend avec l'ajout du site et du projet de recherche grippenet.fr, faisant appel au grand public pour la surveillance de la grippe, et avec la comparaison entre les estimations du réseau et le nombre de recherches sur Google concernant ces maladies.

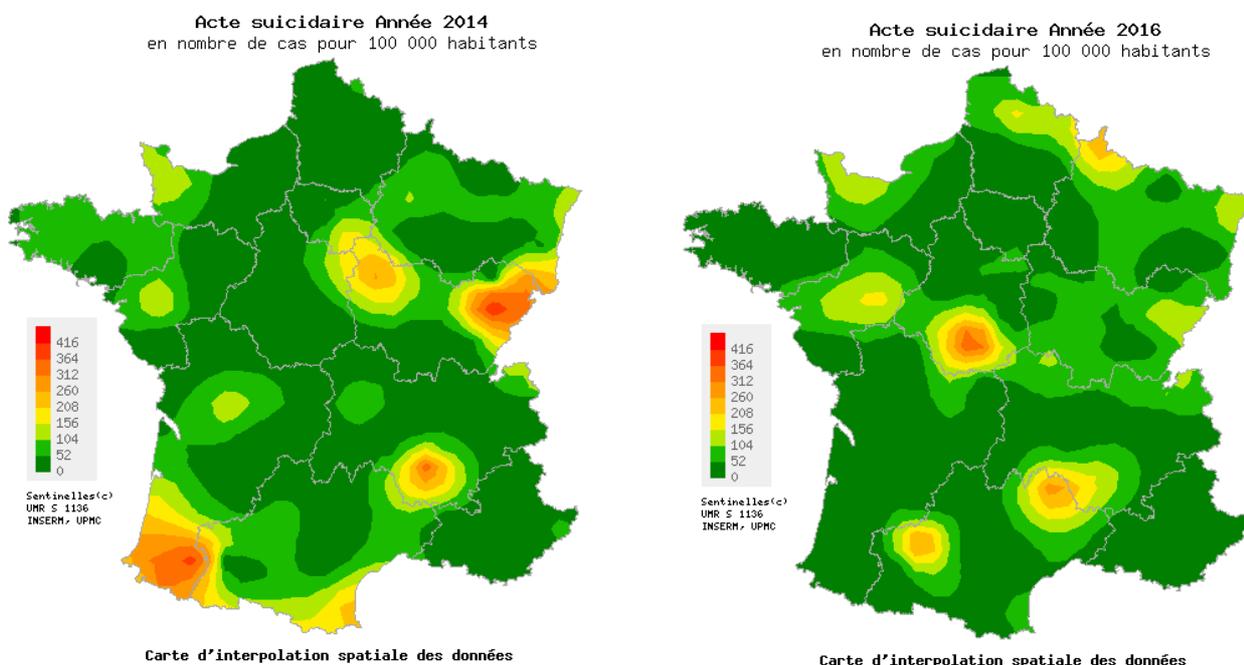


Figure 8 : Surveillance Sentinelle des suicides en France

Cartes de l'incidence de suicides provenant des données du réseau Sentinelle (« Réseau Sentinelles > France > Accueil », s. d.)

L'avantage des données de santé publique par rapport à celles précédemment décrites tient à leur fiabilité et leur maniabilité. En revanche, elles se présentent souvent sous une forme agrégée au niveau d'une population, et non comme données individualisées, limitant les recoupements possibles et l'intérêt de leurs conclusions pour une personne donnée. Toujours est-il qu'elles illustrent bien le fait que la définition de Big Data tient autant aux outils et au type de traitement des données, qu'à la nature-même des données.

3. Outils d'exploitation des données massives

Les exemples précédents donnent un aperçu des problématiques techniques soulevées par les données massives, particulièrement en médecine, et font allusion à un ensemble de techniques développées pour les résoudre. Nous allons maintenant explorer, avec un peu plus de détail, cette boîte à outils pour mieux comprendre ce qui se passe dans le traitement des données.

a. Infrastructure de stockage et de traitement²

Les ordinateurs et disques durs actuels, aux alentours de 250 gigaoctets (250×10^9 octets) à 1 téraoctet (10^{12} octets), possèdent une mémoire généralement suffisante pour un usage personnel, même avec l'avalanche de données que l'on génère ou que l'on télécharge. Cependant, les données de santé décrites précédemment représentent un volume nettement plus important. À titre d'exemple, le séquençage du génome entier d'une seule personne peut occuper un fichier d'une centaine de gigaoctets. Certains auteurs estiment que les données génomiques à elles seules dépasseront les besoins en archivage et en traitement de YouTube, Twitter, et de la recherche en astronomie (Stephens et al., 2015). Même avec la diminution du coût de stockage, la gestion de bases de données massives nécessite davantage de ressources que la simple accumulation d'un grand nombre de disques durs. En effet, pour espérer pleinement les exploiter, ces bases doivent avoir une mémoire de stockage adéquate, être facilement accessibles depuis des lieux distincts, par de nombreux utilisateurs simultanés, et permettre une puissance de traitement suffisante pour des opérations complexes en temps réel (Chervenak, Foster, Kesselman, Salisbury, & Tuecke, 2000).

Les super-ordinateurs remplissent ces exigences, et les capacités chez Watson Health (IBM), ainsi que le système BioWulf au NIH sont des exemples d'utilisation

² En plus des références citées, différentes sources destinées aux développeurs et programmeurs, comme le site stackoverflow.com (« Stack Overflow - Where Developers Learn, Share, & Build Careers », s. d.), et des discussions personnelles ont permis de vérifier, d'enrichir ou de mettre à jour les informations obtenues dans les publications citées, pour élaborer le contenu du sous-chapitre a. et en partie b.

biomédicale de tels systèmes. Mais les coûts de développement ou les modalités d'accès à un super-ordinateur ne sont pas adaptés à tous les besoins ou à tous les établissements. En revanche, des systèmes plus modulables existent, qui reposent sur des utilisations innovantes d'équipement existant, et sont basés sur différentes manières d'interconnecter un réseau de processeurs et de mémoires de stockage. Grâce à des processeurs et des mémoires reliés en réseau, et une structure logicielle (*framework*) conçue pour les coordonner, il est possible de construire un système de traitement et de stockage distribué ou parallèle, où les fichiers sont divisés en différents fragments et distribués parmi les différentes mémoires de stockage, et où les opérations de calcul sont divisées en plusieurs sous-tâches pour chaque processeur. De cette manière, on peut utiliser du matériel plus couramment disponible pour chaque composant individuel du système, tout en bénéficiant d'une capacité de stockage et de traitement proche ou équivalente à un super-ordinateur. Le système étant évolutif, on peut utiliser uniquement la capacité nécessaire pour une tâche donnée, ou ajouter des composants supplémentaires en fonction de l'évolution des besoins.

En pratique, ce système de traitement en parallèle ou de traitement distribué (*parallel/distributed computing*) se décline sous plusieurs formats, qui diffèrent selon le degré de synchronisation, de dépendance et d'utilisation des processeurs (Luo et al., 2016). On peut distinguer de manière grossière :

- microprocesseurs multi-coeurs (*multicore computing*) : en quelque sorte le niveau le plus élémentaire, il s'agit de l'intégration de plusieurs unités de calcul (les coeurs) au sein d'une même puce ; cette technologie existe déjà couramment au sein des ordinateurs personnels,
- grille informatique (*grid computing*) : des ressources informatiques hétérogènes, délocalisées et autonomes (ordinateurs personnels, smartphones, serveurs etc) sont liées entre elles et forment une infrastructure virtuelle distribuée, partageant chacune leurs fonctionnalités de manière autonome, fournissant ainsi une capacité de stockage et de traitement large mais variable, non-synchronisée,
- grappe de serveurs (*cluster computing*) : un ensemble de processeurs homogènes, les « noeuds » (*nodes*), sont connectés par un réseau local et chaque « noeud » dans le réseau accomplit la même tâche, contrairement à l'organisation en grille où les machines sont autonomes ; cette configuration est adaptée aux tâches nécessitant à la fois un grand volume de données et une grande synchronicité (bande passante élevée, faible latence : *high bandwidth, low latency*),

- traitement massivement parallèle (*massively parallel computing*) : ce cas de figure peut être considéré comme un cas particulier de l'organisation en grappe, avec un très grand nombre de processeurs mis en lien par des connexions spécialisées,
- *general-purpose computing on graphics processing units* (GPGPU) : emblématique de la culture de « bricolage » dans le milieu informatique, il s'agit de l'utilisation, pour un traitement informatique générique, de cartes graphiques grand public, qui étaient conçues à l'origine pour le traitement graphique intensif nécessaire pour les jeux vidéos. Les cartes graphiques ont déjà une organisation et un traitement de données en parallèle, peuvent être associées à un processeur central ou additionnées pour augmenter la capacité de traitement.

L'architecture en parallèle ou distribuée est particulièrement adaptée au stockage et au traitement caractérisés par des tâches mathématiquement intenses et répétitives. Pour la mettre en œuvre, des structures logicielles (*framework*) adaptées ont été développées, comme MapReduce par Google, ou les solutions open-source Hadoop et Spark développées par Apache Software Foundation (Luo et al., 2016). Leur fonctionnement repose sur la fragmentation et la distribution des fichiers au sein des différents nœuds du système (*distributed file system*), puis la fragmentation des tâches parmi ces nœuds, de manière à ce que chaque nœud traite les données de manière localisée, évitant ainsi des transferts itératifs de données chronophages. Les différents traitements sont ensuite regroupés et synchronisés. Le système peut ainsi fonctionner sans matériel de connexion spécialisé, utilisant uniquement du matériel communément disponible et parfois déjà installé.

Ces ressources sont accessibles aux utilisateurs sous différentes modalités. Les établissements avec suffisamment de ressources ont leurs propres centres de données (*data center*) et serveurs qui peuvent être conçus selon une architecture d'informatique parallèle/distribuée. Cela nécessite non seulement un local dédié pour l'équipement informatique, avec des conditions climatiques, de sécurité et d'alimentation adaptées, mais aussi un entretien physique et logiciel régulier. Cette modalité offre un niveau de contrôle et de lisibilité maximal, mais à un certain coût. Plus récemment, dans une mouvance considérant les ressources informatiques davantage comme une infrastructure de services qu'une ressource matérielle, les modalités d'usage et les objectifs commerciaux tendent à se focaliser sur l'accès aux fonctionnalités du matériel informatique (plutôt que l'accès au matériel en lui-même). C'est ainsi que se développe le *cloud computing*, qui délocalise et centralise l'infrastructure matérielle, pour pouvoir

faciliter l'accès à des fonctionnalités variées à un coût réduit, grâce à l'économie d'échelle. Les utilisateurs évitent ainsi l'investissement monétaire, la responsabilité et la rigidité associés à la possession d'un serveur fixe, mais perdent en contrepartie, un degré d'autonomie et de contrôle sur le système qui est partagé avec d'autres utilisateurs. Typiquement, les clients louent des services, comme le stockage de données (mesuré en octets), la puissance de traitement (location à l'heure, ou forfait, différentes capacités de traitement), en fonction de leurs besoins, et la connexion avec le fournisseur de services se fait par internet. De nombreux acteurs offrent des services de ce type, allant des géants du numérique (Amazon Web Services, Microsoft Azure, Oracle etc) à des initiatives de centres de recherche comme BOINC (Berkeley Open Infrastructure for Network Computing).

La sécurité, le risque de perte des données et les dysfonctionnements d'équipement sont en jeu dans tout système d'accès à distance des données, que ce soit sur le cloud ou sur un serveur privé. La plus grande visibilité des fournisseurs de services *cloud* en font peut-être davantage une cible de cyberattaques que des serveurs locaux ou privés. Mais ils ont aussi plus d'intérêt et de ressources à dédier à l'entretien et à la sécurité de leurs systèmes (et les serveurs locaux ne sont en tous cas pas épargnés par ce genre d'attaques). En revanche, le recours à un fournisseur de services, un tiers, introduit un degré d'opacité quant aux questions de propriété des données et de responsabilité juridique, compte tenu des différences législatives entre pays, de la variabilité de conditions de service et de contrats client.

b. Traitement algorithmique et techniques de visualisation des données

Une manière de comprendre le traitement de données serait de le diviser en trois tâches :

- récupérer et compiler les données : extraction de données et d'information,
- chercher un sens aux données : analyse statistique ou algorithmique,
- présenter et communiquer cette nouvelle information ou connaissance : visualisation des données.

Extraction de données

Les données que l'on souhaite analyser ne se présentent pas toujours sous la forme d'informations directement exploitables, comme un tableau de valeurs dans un fichier Excel. Or il n'est pas envisageable de parcourir manuellement des bases de données énormes pour en extraire les informations, d'où l'apparition d'outils automatisant cette tâche. Il s'agit le plus souvent de logiciels entiers ou de séquences de programmes, soit prêts à utiliser, soit modifiables, ou créés ad hoc, chargés de reconnaître les informations recherchées, de les copier et de les compiler sous un format adapté. Cette activité est connue sous le nom de *data scraping* (râclage de données) ou *information extraction*. Elle diffère de l'utilisation d'un moteur de recherche, car l'objectif n'est pas de trouver une référence, mais de capter et d'organiser le contenu d'une référence, qu'il s'agisse d'un site web, d'un document ou autre. Le fonctionnement précis de ces outils ne sera pas détaillé ici, et est très variable car il dépend de l'information à extraire et du format du support original : certains outils s'appuient sur l'architecture d'un site web, d'autres capturent des informations textuelles à partir de clichés grâce à la reconnaissance optique de caractères (OCR), d'autres encore analysent un texte libre etc.

Cette dernière catégorie a un intérêt particulier, car elle correspond aux techniques de traitement automatique du langage naturel (TALN ou *natural-language processing*, NLP, en anglais), indispensables pour l'exploitation automatisée des dossiers médicaux. Meystre et al. proposent une revue et une critique des outils développés pour l'extraction d'informations à partir des dossiers médicaux numérisés, qu'il s'agisse de domaines structurés ou de texte libre (Meystre et al., 2008). Ils présentent une variété d'approches, allant de la reconnaissance de motifs stéréotypés (pour détecter, par exemple le format standardisé d'un numéro de sécurité sociale), aux algorithmes basés sur des règles sémantiques et syntaxiques, à l'apprentissage automatique (*machine learning*) par des méthodes statistiques. Les auteurs décrivent quelques utilisations comme l'extraction de codages, de libellés diagnostiques ou sémiologiques, l'analyse contextuelle (indicateurs temporels, négation), l'établissement de lexiques médicaux, l'aide à la décision médicale ou la surveillance. Bien que plusieurs logiciels aient démontré une bonne efficacité dans le traitement de dossiers médicaux, notamment en apprentissage automatique statistique, les auteurs mettent en relief la difficulté à développer ces outils, nécessitant une intervention humaine significative dans la programmation des logiciels ou dans l'élaboration de larges bases de données annotées pour les algorithmes d'apprentissage

automatique. Ils mettent en garde contre les limites des études de validation de ces logiciels, et les limites à leur généralisation hors de leur contexte de développement. Leur conclusion rappelle que les textes cliniques sont simplement plus difficiles à analyser que d'autres types de textes.

Un aspect important de l'extraction des données est leur standardisation, afin de faciliter leur partage, leur comparabilité, leur lecture et leur exploitation par un même algorithme. En imagerie, par exemple, de nombreuses étapes de traitement d'image (suppression du bruit et des artefacts d'appareil, correction des artefacts de mouvement, transformation et alignement à un espace de référence) sont nécessaires pour optimiser le signal qui sera utilisé comme donnée d'entrée dans les algorithmes d'apprentissage automatique (voir par exemple, la description détaillée des méthodes dans l'étude de Drysdale et al., 2017). De même, en génétique, la détection d'erreurs de séquençage, la mise en correspondance des données de séquençage avec des génomes de référence, et la standardisation de données provenant de différents logiciels de séquençage, sont toutes des étapes nécessaires – et souvent laborieuses, même avec de nombreux outils pouvant automatiser ces tâches – avant de pouvoir mettre ces données à profit dans une analyse de variations génotypiques (Luo et al., 2016; Sullivan, 2010). Un adage fréquent chez les *data engineers* et *data scientists* dit que leur travail est 80 % de nettoyage de données et 20 % d'analyse (J. Anderson, 2018; Lohr, 2014), ce qui est d'autant plus vrai en biomédecine avec l'exigence de données propres de haute qualité, très éloignée de la réalité de données fragmentées, non standardisées, incomplètes et de qualité variable.

Visualisation de données

L'analyse de données doit servir à mieux comprendre les données et les relations entre elles. La manière dont cette analyse est présentée visuellement peut contribuer à rendre les données plus intelligibles et plus éclairantes. La présentation visuelle a aussi pour objectif la communication de cette analyse à un public. Une attention particulière sera donc apportée au choix des outils et des modalités de visualisation qui dépendront des données, du public et surtout de la problématique à laquelle l'analyse apporte une réponse. L'émergence, dans les années 90, de la *data visualisation* (visualisation de données) comme spécialité dans le domaine du graphisme informatique (*computer graphics*), reflète l'importance de cette étape dans le traitement des données, notamment les données complexes et multimodales (Post, Nielson, & Bonneau, 2003). Au carrefour

entre informatique, design et communication, la visualisation de données peut aussi influencer l'étape d'analyse de données, bien qu'à première vue, elle ne concerne que le produit final du traitement de données. En fait, on peut être amené à choisir une méthode statistique plutôt qu'une autre, en fonction des modalités de visualisation qu'elle offre ou de ses propriétés graphiques (par exemple, un arbre de décision plutôt qu'une régression logistique, James et al., 2013). Les logiciels de bureautique les plus répandus (Microsoft Office et l'équivalent open-source Open Office) ont une offre de techniques de visualisation assez bien développée pour une utilisation courante, et simples à mettre en œuvre. Les logiciels et environnements de programmation MATLAB et R (logiciel libre) ont été utilisés depuis longtemps en recherche pour le traitement et la visualisation de données. D'autres langages et environnement de programmation les ont rejoint, comme Scilab, logiciel libre se basant sur MATLAB, et les langages Python et Java, pour n'en citer que quelques uns. Cette diversité de logiciels et de langages de programmation offre une variété de modalités de présentation des plus générales aux plus personnalisables, des plus simples d'utilisation à celles requérant un niveau d'expertise en programmation. On retrouve :

- les diagrammes classiques comme les diagrammes en bâton, histogrammes, diagrammes circulaires et nuages de points,
- les *heatmaps* (cartes de chaleur) pour une visualisation rapide de quantités ou d'intensités par le degré de variations de couleurs,
- les représentations cartographiques pour des valeurs géolocalisées,
- les boîtes à moustache pour une représentation des moyennes, médianes et variations,
- la visualisation de réseaux faisant apparaître l'importance des nœuds, et l'intensité des relations entre les différents éléments d'un réseau,
- les dendrogrammes pour le regroupement de données hiérarchisées,
- la possibilité d'améliorer toutes ces représentations par des animations permettant de visualiser des aspects dynamiques ou évolutifs, et par des fonctionnalités interactives faisant apparaître des données plus détaillées,
- de nombreuses autres formes non-classifiables, créées pour être à la fois informatives et esthétiques.

Analyse de données³

Que l'on ait compilé une base de données de manière systématique, ou avec l'objectif d'une étude précise, on cherche à *comprendre* ces données, c'est-à-dire en saisir le sens, et savoir quoi et comment changer à la lumière de cette nouvelle connaissance. On peut voir cela, en somme, comme un problème d'apprentissage à son niveau le plus élémentaire : formuler une représentation généralisée de propriétés ou d'un fonctionnement à partir de quelques observations, cette représentation pouvant ensuite guider les réponses ultérieures à de nouvelles observations. L'externalisation de cette faculté à un système informatique, donne l'apprentissage automatique (*machine learning*), une spécialisation des sciences informatiques et des sciences des données en plein essor.

Certaines techniques proposées par le passé pour l'apprentissage automatique se basaient sur la programmation de règles de déduction logique ou des systèmes experts. Mais les techniques qui se répandent actuellement sont fondées sur les statistiques et les probabilités, se rapprochant peut-être davantage d'un mode d'apprentissage humain, inférentiel et probabiliste (Tenenbaum, Griffiths, & Kemp, 2006). Le problème devient alors : soit une base de données avec n échantillons, caractérisés par un nombre p de variables⁴ $X (X_1, X_2 \dots X_p)$. On peut se demander quelle est la relation entre les différents échantillons en fonction des variables X : s'organisent-ils en groupes identifiables ? On peut aussi chercher à comprendre la relation qui lie ces n échantillons à des réponses Y qui leur sont associées, ce qui revient à chercher la fonction $f(X)$ telle que : $Y = f(X) + \epsilon$ (où ϵ est l'erreur irréductible indépendante de X). Le travail de l'apprentissage statistique, dans ce cas, se résume donc à estimer une fonction $f_{est}(X)$ qui se rapproche le plus possible de $f(X)$ (James et al., 2013).

Les deux questions précédentes font référence à deux catégories distinctes d'apprentissage automatique : l'apprentissage supervisé et non-supervisé. En apprentissage supervisé où les échantillons sont associés à des résultats Y , les outils statistiques répondent aux questions :

3 Les informations contenues dans ce sous-chapitre sont largement basées sur et adaptées des manuels de référence *An Introduction to Statistical Learning* (James, Witten, Hastie, & Tibshirani, 2013) et *The Elements of Statistical Learning* (Hastie, Tibshirani, & Friedman, 2017).

4 Aussi appelée variable indépendante, Y étant la variable dépendante. Nous utiliserons ici la dénomination « variable » et « résultat » pour X et Y respectivement, pour être plus proches de « variable » et « output » dans le texte de James et al., avec l'avantage supplémentaire de moins alourdir le texte.

- d'inférence : comment se comporte le résultat Y quand les variables X changent, c'est-à-dire, quelles variables X sont associées à la réponse Y , et quelle type de relation lie ces variables à Y ? Ici, on est intéressé par la forme globale de la fonction f .
- de prédiction : pour un tel X donné, quel est le résultat Y attendu ? On est moins préoccupé par la forme de f que par la justesse du résultat prédit localement.

Le terme « supervisé » décrit le fait que l'on fournit à la machine le résultat attendu, tel un enseignant qui indique à un enfant l'objectif à atteindre, alors que dans l'enseignement non-supervisé, sans résultat Y , il n'y a pas d'objectif défini, pas de fonction $f(X)$ à estimer. On ne cherche pas à prédire un résultat, mais à découvrir des motifs, des relations ou des sous-groupes parmi les échantillons ou les variables X , comme lorsqu'un enfant, face à un mélange de jetons bleus et rouges, ronds et carrés, trouve de lui-même un principe pour les organiser et les départager. L'on pourrait peut-être dire que l'apprentissage supervisé sert à mieux *expliquer*, alors que l'apprentissage non-supervisé sert à mieux *décrire*. À la différence des statistiques « classiques » dont le but essentiel, en sciences bio-médicales, est de déterminer si les variations que l'on observe dans une série de données sont le fruit du hasard, l'apprentissage statistique cherche, de surcroît, à modéliser ces variations, à les définir mathématiquement, pour pouvoir ensuite formuler des prédictions individualisées.

Comment arrive-t-on en pratique à cette modélisation en apprentissage supervisé ? Deux approches globales sont possibles. Dans l'approche paramétrique, on suppose a priori une forme pour f , la forme la plus simple étant la régression linéaire simple dans le cas d'une variable X simple : $Y = f(X) + \epsilon = \beta_0 + \beta_1 X + \epsilon$.

Plutôt que d'ajuster une fonction entière inconnue aux n observations qui constituent les données d'entraînement, le problème est réduit à l'estimation et l'ajustement des coefficients, ou paramètres, β . Mais en général, on ne connaît pas au préalable la forme de la relation entre X et Y – sinon il n'y aurait pas besoin de l'estimer ! Donc avec une approche paramétrique, on court le risque de choisir un modèle qui s'éloigne beaucoup de la réelle forme de f et qui ne s'ajustera pas bien aux données.

L'approche non-paramétrique n'a pas de présupposés sur la forme de f . La définition et l'ajustement de la fonction se fait localement, en construisant selon certaines règles la fonction la plus lisse possible, s'approchant au mieux des données d'entraînement. La fonction sera bien ajustée aux n données à disposition, mais ne sera représentative de la réelle relation f que s'il y a suffisamment de données pour une estimation juste, problème schématisé sur la figure 9 (cette figure ne représente pas le cas

d'un modèle non-paramétrique mais elle illustre le problème du manque d'échantillons suffisants pour l'estimation d'un modèle). Les techniques non-paramétriques nécessitent donc une grande quantité d'observations, une large base de données d'entraînement, pour une estimation valable de f .

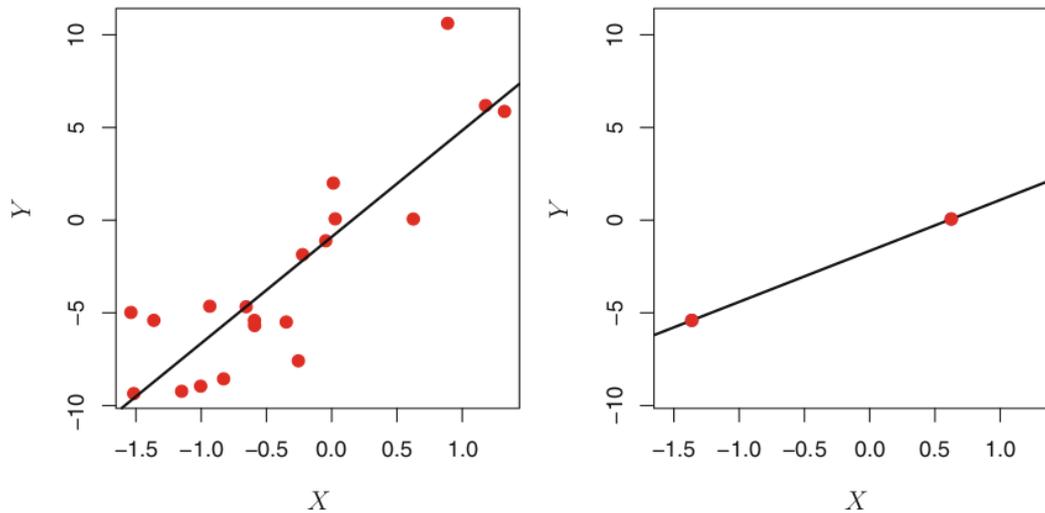


Figure 9 : Nombre n d'échantillons par rapport au nombre de p paramètres

En situation de faible dimension ($n > p$), la fonction peut être mieux estimée (gauche), alors que pour le même jeu de données, elle est mal estimée en cas de n insuffisants par rapport aux paramètres (droite, James, Witten, Hastie, & Tibshirani, 2013)

Cette propriété des outils non-paramétriques évoque la question du surapprentissage (*overfitting* en Anglais). Une fonction parfaitement ajustée aux données d'entraînement est-elle souhaitable ? Non, car une telle fonction risque de ne pas être bien généralisable : trop « collée » aux données d'entraînement, elle sera moins performante face à de nouvelles données qui n'auront pas servi à l'apprentissage. Le risque de surapprentissage augmente avec la flexibilité de la méthode d'apprentissage, notamment celles qui, comme les techniques non-paramétriques, n'ont pas de forme fixe a priori. En effet, une droite résultant d'une fonction linéaire est moins susceptible de « coller » à un nuage de points, qu'une courbe très ondulée résultant d'une méthode non-paramétrique.

Il est donc apparent qu'une estimation utile de f repose sur un équilibre entre l'ajustement aux données d'entraînement et la possibilité d'une généralisation à des données externes. Cet équilibre correspond à la tension entre deux propriétés contradictoires des méthodes d'apprentissage : le compromis biais-variance. Sans détailler les définitions mathématiques, la variance représente une mesure quantitative du changement de f_{est} quand les données d'entraînement changent ; elle reflète, en quelque sorte, la stabilité de f_{est} . Une méthode d'apprentissage plus flexible donnera un résultat

plus instable, variant davantage selon les modifications des données d'entraînement : elle aura donc souvent une plus grande variance. Le biais représente l'erreur introduite par le choix d'un modèle trop simplifié par rapport à la réelle fonction f entre X et Y . Par exemple, si on choisit une méthode linéaire, alors que la vraie relation entre variables et réponses dévie beaucoup de la linéarité, la fonction estimée f_{est} aura un biais élevé, même si on essaie de mieux l'estimer en élargissant le jeu de données d'entraînement, car le modèle est trop éloigné de la réalité. Plus une méthode d'apprentissage est flexible, en général, moins elle sera biaisée. L'estimation optimale doit avoir une faible variance et un biais faible, mais on voit que ces deux propriétés évoluent dans des directions opposées.

Comment arrive-t-on à un compromis acceptable ? Cette question nous amène aux techniques d'ajustement et de validation. On ajuste f_{est} en essayant de minimiser l'écart entre le résultat estimé $Y_{est} = f_{est}(X) + \epsilon$ et le résultat attendu Y observé dans les données d'entraînement. En régression linéaire, cela revient à minimiser la somme résiduelle des carrés $(Y - Y_{est})^2$ sur les données d'entraînement (RSS ou *residual sum of squares*), somme que l'on retrouve aussi dans le calcul de l'erreur quadratique moyenne (*mean squared error* ou MSE), et du coefficient de détermination R^2 (qui est une indication de la proportion de variance expliquée par le modèle). Mais le niveau d'erreur sur les données d'entraînement n'est pas un bon indicateur du niveau d'erreur que la fonction aura face à de nouvelles données, comme indiqué par la figure 10. En effet, on peut avoir sur-entraîné le modèle, auquel cas l'erreur sur les données d'entraînement sera très faible, mais elle sera très élevée sur des données externes. Pour cette raison, une estimation ne sera pas valable tant qu'elle n'aura pas été validée sur un jeu de données test, différent des données d'entraînement. De même, la performance d'une estimation sera donnée par le niveau d'erreur test (*test MSE*) et non le niveau d'erreur d'entraînement. La méthode de référence pour l'étape de validation est la vérification de la fonction estimée sur un jeu de données test externe, totalement indépendant des données d'entraînement. Mais si une telle série de données n'est pas disponible, il existe d'autres méthodes se basant sur les données d'entraînement (*cross-validation, bootstrap*). Pour revenir à la question initiale du compromis biais-variance, le lien entre l'erreur test et ce compromis paraît donc intuitif : une fonction estimée aura une erreur test faible seulement si elle est suffisamment flexible pour bien rendre compte des données (biais faible), tout en étant assez stable et généralisable pour être appliquée à des données externes (variance faible). En fait, ce lien intuitif a une traduction mathématique, car l'erreur test moyenne sur de multiples

estimations de f_{est} pour un x_0 donné n'est autre que la somme de trois valeurs : la variance de f_{est} en x_0 , le biais au carré de f_{est} en x_0 et la variance de l'erreur irréductible ϵ .

$$\text{Erreur test moyenne} = \text{Var}(f_{est}(x_0)) + [\text{Biais}(f_{est}(x_0))]^2 + \text{Var}(\epsilon)$$

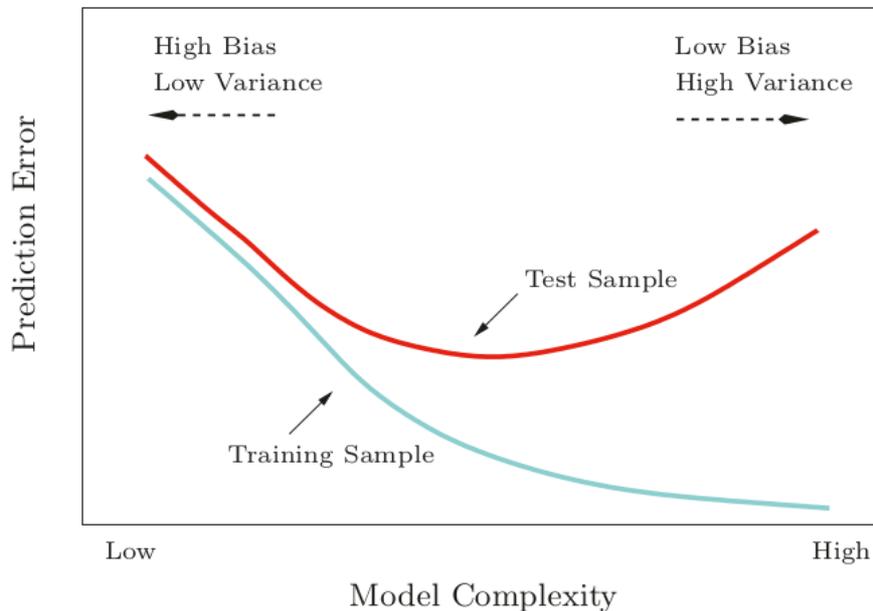


Figure 10 : Compromis biais-variance et différence entre erreur d'entraînement et erreur test (Hastie, Tibshirani, & Friedman, 2017)

En résumé, ces deux étapes d'ajustement et de validation externe permettent de confronter le modèle aux deux problèmes de biais et de variance, et de trouver un compromis acceptable – ou le cas échéant, de se rendre compte qu'il faut choisir un modèle différent.

Quels sont plus précisément ces modèles, paramétriques ou non-paramétriques ? Nous avons vu précédemment que la régression linéaire était l'exemple le plus simple de modèle paramétrique, et historiquement le mieux connu. Elle est relativement facile à interpréter grâce aux coefficients β , qui donnent une représentation quantitative de la relation entre variable X et réponse Y . De plus, de nombreux outils d'inférence (tester l'hypothèse nulle en calculant l'intervalle de confiance et la valeur-p) sont facilement et largement utilisables en régression linéaire, permettant d'évaluer si la relation observée entre X et Y est due au hasard ou non. On peut même représenter des variables multiples, par une régression linéaire multiple qui prend la forme additive :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon.$$

Or la nature des données complexes et multiformes caractérisant le Big Data, ne se prête pas toujours à un modèle aussi rigide, induisant un biais important. Le défi posé par ces données tient principalement au fait qu'elles peuvent être :

- hautement multivariées : un grand nombre de variables X liées à une réponse Y ,
- corrélées entre elles : la multi-colinéarité met à mal le présupposé d'indépendance des variables nécessaire à la régression linéaire simple ou multiple,
- liées par une relation non-linéaire.

On peut gérer cette complexité, d'un côté, en introduisant des degrés de flexibilité dans les méthodes d'apprentissage paramétriques pour estimer des fonctions non-linéaires. En effet, de nombreuses méthodes utilisées actuellement en apprentissage automatique sont des généralisations ou des extensions de la régression linéaire.

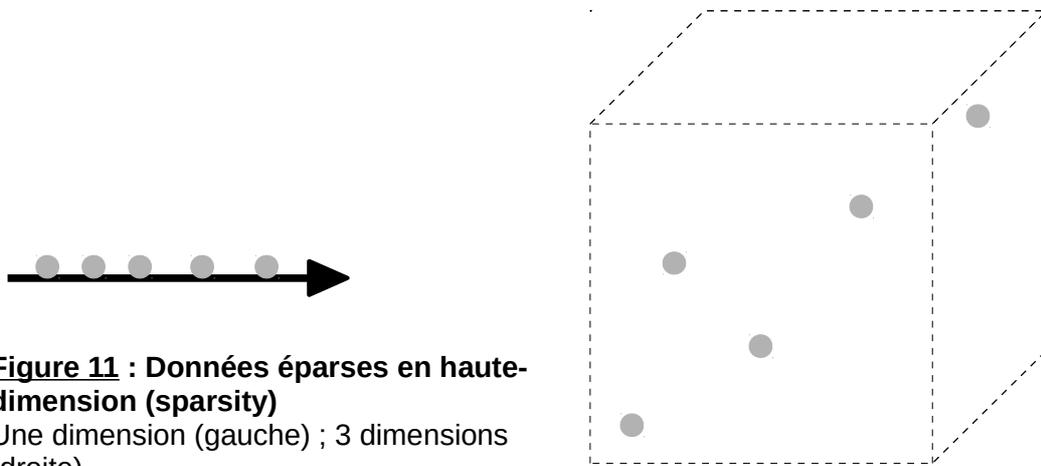


Figure 11 : Données éparées en haute-dimension (sparsity)
Une dimension (gauche) ; 3 dimensions (droite)

Mais le problème de la haute-dimensionalité est bien plus épineux, à tel point qu'il est nommé « curse of dimensionality » (fléau de la dimension) (James et al., 2013). Lorsque le nombre de dimensions augmente, les échantillons sont plus « éparpillés » au sein des différentes dimensions : on dit que les données sont plus éparées, qu'il y a une plus faible densité des données. On peut visualiser cette notion très simplement en imaginant dix billes dispersées sur une ligne (une dimension), puis ces même dix billes dispersées dans un cube (trois dimensions, Figure 11). Les billes seront plus éloignées les unes des autres dans le cube par rapport à la ligne, elles seront moins densément regroupées. Ceci aura deux conséquences : en augmentant les dimensions, la variance de la fonction estimée augmentera énormément, et avec, le risque de surapprentissage, et l'erreur test. Cet effet est aussi valable pour les méthodes non-paramétriques qui peuvent être de base plus flexibles et variables. De plus, l'étendue de la multi-colinéarité en haute-dimension (nombre p de variables $>$ nombre n d'échantillons), empêche de savoir quelles variables sont réellement prédictives de la réponse. La plupart des outils de statistiques

classiques sont adaptés aux données à faible dimension, et perdent en fiabilité face à la haute-dimension : valeur-p moins interprétable, test de l'hypothèse nulle moins puissant, diminution de l'erreur d'entraînement et augmentation du R^2 d'entraînement avec l'ajout de nouvelles variables, indépendamment de leur réel effet prédictif.

Ceci montre que le problème des données complexes n'est pas seulement une question de non-linéarité. Il sape notre capacité d'inférence autant en apprentissage automatique qu'en apprentissage « humain ». On tente donc, d'un autre côté, de réduire cette complexité en se basant directement sur les données, en utilisant une manière statistique, qui ne soit ni arbitraire (limiter les variables un peu au hasard, ou post hoc), ni ne dépende d'une expertise dans le domaine étudié (limiter les variables en se basant sur des hypothèses issues de la littérature ou d'études précédentes). Cette étape correspond aux techniques automatiques de sélection, de réduction, de régularisation des variables/dimensions.

Par conséquent, la construction du modèle, de l'estimation de $f(X)$, procède d'une part en accommodant la complexité des données par un modèle plus flexible, et d'autre part en posant des contraintes à la complexité. Les techniques suivantes sont quelques unes des principales utilisées pour atteindre ces objectifs, et sont également résumées dans le tableau 1 :

- gérer la non-linéarité :
 - x régression polynomiale : $X_1^2, X_1^3 \dots$
 - x approches non-paramétriques, ajustant localement une fonction constante, une régression linéaire ou une régression polynomiale, sur des intervalles d'une variable X :
 - fonctions en escalier (*step functions, piecewise constant functions*)
 - régression par splines (*regression splines*)
 - *smoothing splines*
 - régression locale (*local regression*)
 - x approches non-paramétriques sans lien avec la régression linéaire :
 - méthode des k plus proches voisins (*k-nearest neighbor regression* ou KNN)
 - arbres décisionnels en régression (*decision trees for regression*), améliorés par les techniques de *bagging*, des forêts d'arbres décisionnels (*random forests*) et de *boosting*
- accommoder de nombreuses variables :

- x régression linéaire multiple : la réponse Y est la somme de régressions linéaires simples pour chaque variable X
- x régression linéaire avec produit de variables : X_1X_2
- x modèles additifs généralisés (*generalized additive models* ou GAM) : la réponse Y est la somme de fonctions f différentes pour chaque variable X
- minimiser la complexité :
 - x par la sélection d'un sous-groupe de variables pertinentes, en testant différentes combinaisons de variables : *subset selection* (*best subset selection*, *stepwise selection*, *backward stepwise* par exemple)
 - x en imposant une pénalité sur les coefficients de régression (*shrinkage*), permettant ainsi de rétrécir l'effet des variables moins pertinentes ou de les annuler (*ridge regression*, *lasso*, *elastic net regularization*)
 - x en projetant les p variables d'origine en un nombre inférieur de nouvelles variables ou de vecteurs, qui combinent l'information contenue dans les variables d'origine : *dimension reduction* (régression sur composantes principales, ou *principal components regression*, et moindres carrés partiels, ou *partial least squares* par exemple).

Ces techniques sont disponibles dans les logiciels de traitement de données, comme R. La sélection d'une méthode plutôt qu'une autre dépend des atouts et des faiblesses de ladite technique face au nombre de variables et de leur degré de corrélation, à la quantité de données d'entraînement, au lien supposé entre variables et réponse, à la question posée (prédiction, inférence) et au degré de précision demandé. En pratique, on mettra en œuvre plusieurs méthodes de sélection de variables et d'apprentissage. On choisira celles qui auront la meilleure performance en validation sur des données test, tout en tenant compte du fait qu'un modèle plus simple est aussi plus facile à comprendre et à expliquer, donc peut être plus adapté aux questions où une justification est nécessaire (inférence, aide au diagnostic et à la décision médicale etc).

Les explications précédentes concernaient implicitement le cas où la réponse est quantitative, où l'apprentissage peut prendre la forme d'une régression de Y sur X . Quand Y est de nature qualitative – on parle de réponse catégorique – l'apprentissage est une tâche de classification pour laquelle plusieurs techniques existent. Certaines sont basées sur des calculs de probabilité, d'autres sur des règles de séparation des variables ou des observations. Leur précision est définie par leur spécificité et leur sensibilité, souvent résumées dans une courbe ROC (*receiver operating characteristic*, tracé de la sensibilité

en fonction de 1-spécificité), l'aire sous cette courbe (*area under curve* ou AUC), ou leur *accuracy* (précision = (vrais positifs + vrais négatifs)/(vrais cas + vrais témoins)). Tout comme en régression, les questions de biais-variance, de surapprentissage et de validation influencent le choix de la méthode d'apprentissage. On trouve, par exemple :

- la régression logistique : estime directement la probabilité de Y sachant X , $P(Y|X)$,
- l'analyse discriminante linéaire, quadratique, multivariée (*linear/quadratic/multivariate discriminant analysis* ou LDA/QDA) : extrapole $P(Y|X)$ du théorème de Bayes après estimation de la probabilité de X sachant Y (la densité conditionnelle de X),
- la classification KNN : méthode non-paramétrique, classifie une observation donnée dans la catégorie correspondant à la majorité des k observations avoisinantes,
- les arbres de décision : non-paramétrique, segmentent l'espace des variables X en sous-régions de telle manière à minimiser l'erreur de classification,
- les machines à vecteurs de support (*support vector machine* ou SVM) : trouvent l'hyperplan séparant géométriquement les n observations en deux catégories dans l'espace des variables, de manière à maximiser la marge entre les deux catégories, c'est-à-dire à être le plus discriminant (les observations à la marge étant des vecteurs qui « soutiennent » l'hyperplan et définissent la marge).

Nous avons donc vu, dans les grandes lignes, les principes de l'apprentissage automatique statistique supervisé. On construit statistiquement un modèle pour pouvoir ensuite prédire un résultat ou inférer une relation. Ce modèle se construit sur des données d'entraînement, mais doit être validé sur des données test ou par validation croisée, et dans l'idéal, être testé sur un échantillon externe complètement différent de l'échantillon d'entraînement et de validation initiale (site de collecte, investigateurs différents par exemple). Le surapprentissage, ou une variance trop importante du modèle, freinent sa généralisation, tandis qu'un modèle trop biaisé ne rend pas bien compte des données. On cherche donc un équilibre entre biais et variance en augmentant un peu la flexibilité du modèle, tout en réduisant la complexité « inutile », toujours avec une validation externe. Il n'existe pas de méthode d'apprentissage supérieure aux autres dans l'absolu, car le choix de la méthode dépend des données et de la question posée. Et on doit être très vigilant aux limites à l'interprétation en haute-dimension.

L'apprentissage non-supervisé présente un autre cadre d'approche des données. Si l'apprentissage supervisé est une dissertation, l'apprentissage non-supervisé est comme un commentaire de texte. On cherche des motifs récurrents dans les variables ou les observations, des manières de visualiser les données. C'est une exploration des données

qui peut générer de nouvelles hypothèses à tester, mais pas de prédictions ni d'inférence. L'abstraction des données peut procéder de plusieurs façons :

- en cherchant une représentation des données à plus faible dimension, qui rend compte au mieux des variations de ces données : en PCA, ceci nous donne un vecteur dans l'espace de variables indiquant la direction de plus grande variation des données⁵,
- en cherchant des similitudes entre les données pour les partitionner (*clustering*) dans des sous-groupe homogènes de variables ou d'observations : algorithmes des *k*-moyennes (*k-means clustering*) ou regroupement hiérarchique (*hierarchical clustering*) par exemple.

Dans l'algorithme des *k*-moyennes, le nombre *k* de clusters (ou grappes) est prédéfini, puis les données sont attribuées à un seul cluster de manière à minimiser la variation intra-cluster. En regroupement hiérarchique, que l'on retrouve par exemple dans des études génétiques, il n'y a pas de nombre fixe de clusters : l'algorithme construit un partitionnement partant des paires d'observations les plus semblables, puis fusionnant les paires les plus semblables, et ainsi de suite jusqu'à l'obtention d'un dendrogramme visualisant des branches imbriquées hiérarchiquement. La hauteur dans le dendrogramme à laquelle deux observations sont fusionnées indique leur degré de ressemblance (plus elle est basse, plus les observations sont proches), et la hauteur à laquelle on choisit de « couper » le dendrogramme définira le nombre de clusters.

L'apprentissage non-supervisé contient une grande part de subjectivité car il n'y a pas de consensus sur la validation des clusters ou vecteurs, et sur l'évaluation de leur significativité. De plus, les choix réalisés dans la mise en œuvre de ces techniques (définition de la similitude et du degré de similitude pour former un cluster, nombre de clusters ou hauteur du dendrogramme, nombre de composants en PCA, étalonnage des données ou non) peuvent donner des résultats très différents, obligeant à effectuer plusieurs essais en variant ces choix, et à estimer la stabilité des clusters obtenus sur des sous-ensembles des données. Surtout, les résultats d'un apprentissage non-supervisé, et notamment d'un partitionnement, ne doivent pas être considérés comme ayant une valeur probante, mais uniquement conjecturale.

5 La PCA est aussi utilisée en apprentissage supervisé pour réduire le nombre de dimensions dans les problèmes à haute dimension (voir p. 46) : il s'agit de *principal components regression*.

Tableau 1 : Résumé des techniques d'apprentissage statistique selon les utilisations

Régression		Classification		
<i>Non-linéarité</i>	régression polynomiale		régression logistique	
	régression locale (local regression)		LDA : analyse discriminante linéaire (linear discriminant analysis), ou quadratique, multivariée...	
	<i>Approches en fonction de base (basis function), avec ajustement local sur intervalles de X</i>	fonctions en escalier : step functions, piecewise constant functions		
		régression par splines (regression splines)		
		smoothing splines		
	<i>Approches non-paramétriques</i>	KNN : méthode des k plus proches voisins (k -nearest neighbor regression)		classification KNN
		Arbres de décision en régression (decision trees)		arbres de décision
		SVM : machine à vecteurs de support (support vector machine)		
<i>Multivariabilité</i>	régression linéaire multiple			
	régression linéaire avec produit de variables (terme d'interaction entre variables)			
	modèles additifs généralisés (generalized additive models)			
<i>Minimiser la complexité</i>	sélection séquentielle de variables (subset selection) : best subset, stepwise selection, backward stepwise...			
	shrinkage/régularisation (terme de pénalité) : ridge regression, lasso, elastic net regularization			
	réduction de dimensions : principal components regression, partial least squares			
Non-supervisée				
Principal component analysis (analyse en composante principale)				
<i>Clustering</i>	k -means clustering (regroupement k -moyennes)			
	hierarchical clustering (regroupement hiérarchique)			

Enfin, on ne peut présenter l'apprentissage automatique sans évoquer l'apprentissage « profond » (*deep learning*), que l'on retrouve derrière plusieurs exemples médiatisés d'intelligence artificielle et qui est parfois confondu à tort avec l'IA dans le discours ambiant, aussi débattue que puisse être la définition de l'IA (Ganascia, 2010)

⁶. L'une des méthodes phare de l'apprentissage profond est celle du réseau de neurones artificiels (*artificial neural network* ou ANN), qui a été appliqué tant dans le diagnostic de rétinopathie diabétique que dans la traduction automatique et le développement

⁶ Mais en fait, toute forme d'apprentissage automatique est une forme d'intelligence artificielle. L'apprentissage automatique est un sous-domaine de l'IA qui, elle, se manifeste aussi sous d'autres formes (Ganascia, 2010).

d'AlphaGo, ordinateur qui a battu les meilleurs joueurs humains au jeu de stratégie Go (Gibney, 2017; Gulshan et al., 2016; LeCun, Bengio, & Hinton, 2015; Singh, Okun, & Jackson, 2017). Cette méthode peut servir à la régression et à la classification (Hastie, Tibshirani, & Friedman, 2017; LeCun et al., 2015). Elle fut développée à la fois dans les domaines des statistiques et de l'intelligence artificielle, faisant écho aux liens, évoqués au début de ce chapitre, entre statistiques et informatique en science des données et en Big Data. Bien que sa conceptualisation date au moins des années 80, avec des éléments datant même des années 60, son essor actuel peut être attribué à son succès dans différents domaines comme la reconnaissance et la classification d'images, grâce à la disponibilité de bases de données massives nécessaires à son entraînement, et à l'augmentation de la puissance de traitement informatique permettant son utilisation plus courante. Son nom reflète son architecture inspirée d'un modèle – certes simplifié – de réseau de neurones, comprenant un stimulus d'entrée (*input*, les variables X) transformé en un signal activant une succession de neurones intermédiaires, pour arriver à une réponse (*output*, réponse Y), et capable d'apprendre comment transformer le signal pour ajuster la réponse.

Son principe est d'une élégante simplicité (Figure 12) : soient des variables X correspondant aux signaux d'entrée (*input*),

1) créer des variables Z (*derived features*) à partir d'une combinaison des variables X dans une fonction d'activation, avec des « poids synaptiques » (*weights*, ou paramètres ou pondérations des variables),

2) réitérer cette transformation dans une ou plusieurs couches successives de « neurones » intermédiaires (*hidden layers* ou *hidden units*) à partir des variables Z de la couche précédente,

3) modéliser la réponse ou cible Y (*output* ou *target*) à partir d'une fonction non-linéaire des dernières variables Z .

Une comparaison avec notre système visuel – la transformation d'un signal lumineux en signal électrique au niveau de la rétine, la notion de potentiel d'action, la propagation de ce signal à travers le nerf optique et sa transformation successive dans les différentes structures sous-corticales et corticales, pour arriver à l'identification d'un objet – saisit bien l'analogie avec le principe des ANN. Il s'agit de *representation learning* : la machine découvre et apprend par elle-même les représentations, c'est-à-dire les attributs ou la combinaison de variables, nécessaires pour trouver la réponse (LeCun et al., 2015). Le schéma cité ci-dessus est le plus élémentaire, mais peut se décliner sous différentes

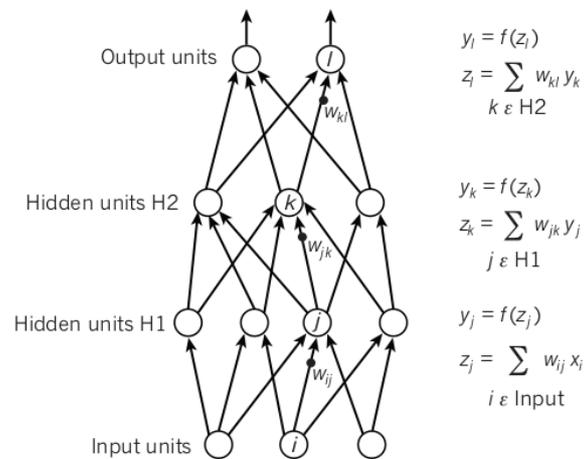
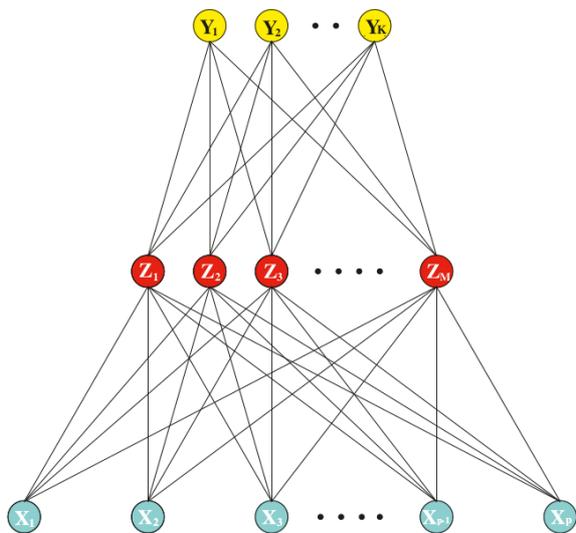
architectures avec des nombres d'unités cachées variables et des connexions circulaires complexes entre les unités cachées. L'architecture choisie dépendra de la nature des données, de leur interdépendance spatiale ou temporelle, de la question posée etc. Cette architecture déterminera une grande part de la performance de l'ANN. Ching et al. offrent une brève description des différentes architectures communément utilisées, et des exemples d'application, que l'on n'abordera pas en détail ici, mais peuvent être consultés en Annexe 3 (Ching et al., 2018).

Le qualificatif d'apprentissage « profond » est lié à la manière dont cet algorithme s'optimise, de manière automatique et itérative à partir des données, par la méthode de la rétropropagation du gradient (*backpropagation*). Rappelons-nous que dans les techniques d'apprentissage citées précédemment, on ajuste des modèles en trouvant les paramètres qui minimisent l'erreur d'entraînement, à savoir la différence entre la réponse Y attendue et la réponse Y estimée, sur *une* fonction donnée, et on sélectionne le modèle dont la forme, les paramètres, les variables, minimisent l'erreur test. Un ANN doit trouver, par l'apprentissage itératif qu'il met en œuvre automatiquement, quelles variables d'entrée prendre en compte dans le signal d'entrée brut, et comment les pondérer pour avoir une fonction qui génère la cible Y . Dans un ANN, les paramètres (*weights*, ou pondérations des unités) sont d'abord assignés plus ou moins au hasard, et après une première transformation du signal d'entrée, on obtient une estimation de Y . La machine calcule alors, comme dans les techniques précédentes, la différence entre la réponse donnée par le modèle et la réponse attendue. Mais cette réponse estimée est le résultat de plusieurs combinaisons de fonctions imbriquées avec un grand nombre de paramètres :

$$Y = f(X) = \sum h(Z) = \sum h(g(X))$$

Or on veut savoir comment ajuster les paramètres dès l'entrée et à travers tout le réseau. Peut-on essayer de calculer directement l'erreur dans chaque unité cachée pour ajuster chaque paramètre localement ? Non, on ne peut pas calculer la différence entre Z estimé et Z attendu dans ces unités cachées car on ne connaît pas le résultat attendu dans ces unités : on ne sait pas ce que doivent être les fonctions $h(Z)$. N'oublions pas que ces unités cachées correspondent à différents modes de représentation des données, générés par l'algorithme, et dont on ne connaît pas d'emblée la signification. Définissons alors la différence entre Y attendu et Y estimé comme une fonction d'erreur. La solution vient d'un peu de mathématiques : le théorème de dérivation des fonctions composées donne le lien entre la dérivée d'une fonction composée et les dérivées des fonctions qui la composent. Quand on dérive la fonction d'erreur, le théorème sus-cité nous donne la dérivée des

fonctions qui la composent, et ainsi, l'erreur au niveau de l'unité cachée précédente. La machine « propage » cette dérivation en arrière, à travers tout le réseau, de la dernière unité à la première, calculant ainsi automatiquement l'erreur à travers tout le réseau. Puis elle ajuste les coefficients selon un taux d'apprentissage que l'on aura défini, et propage les fonctions ajustées en avant, de l'entrée à la sortie. Ce processus est réitéré jusqu'à ce que l'on atteigne un niveau d'erreur satisfaisant. C'est ainsi que la machine « apprend » les représentations qui s'ajustent au mieux aux données : elle ajuste et sélectionne en même temps, « directement » à partir des données brutes fournies.



d Compare outputs with correct answer to get error derivatives

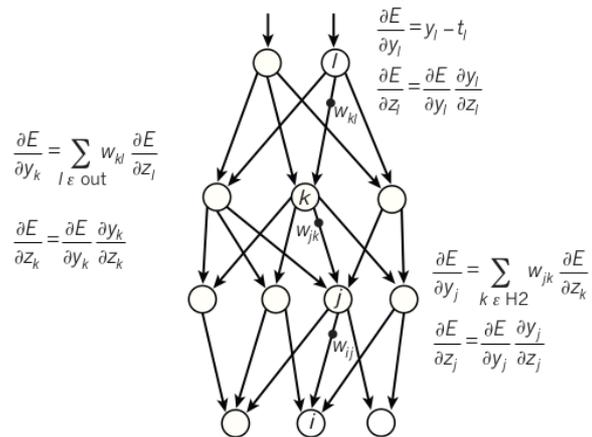


Figure 12 : Réseau de neurones artificiels
Gauche : architecture de base, avec *inputs*, couche cachée simple et *outputs* (Hastie et al., 2017) ; Droite : fonctionnement de base avec *feed-forward* de la fonction (haut), puis rétropropagation du gradient (bas, LeCun, Bengio, & Hinton, 2015)

Bien que simples de principe, les ANN ne sont pas simples d'utilisation, et ne sont pas complètement automatisés. De nombreux choix doivent être faits au préalable qui auront un impact important sur les résultats obtenus, comme l'architecture du réseau, le nombre d'unités cachées, la vitesse d'apprentissage, les pondérations de départ, l'arrêt de l'apprentissage. Comme dans les techniques précédentes, l'inflation du nombre de

variables peut mener au surapprentissage, avec une variance importante et une généralisation faible, d'où la nécessité, encore une fois, d'une validation externe. Mais l'inconvénient majeur de cette technique, surtout en médecine, tient paradoxalement à sa complexité. Malgré leur efficacité en prédiction, les ANN posent des problèmes en inférence. En effet, malgré des techniques récentes en cours de développement (Ribeiro, Singh, & Guestrin, 2016a), il est souvent difficile d'interpréter le modèle de manière à comprendre la nature des liens entre les résultats et les variables, d'expliquer quelles variables influencent le résultat et de quelle façon, et comment le modèle a fait son choix, car les variables d'entrée sont incorporées dans le modèle à des niveaux multiples, subissant des transformations non-linéaires, de manière itérative, augmentant ainsi la complexité de manière exponentielle : c'est le problème de la lisibilité ou de la transparence du modèle.

Ce chapitre est loin d'être exhaustif en ce qui concerne les différentes techniques d'apprentissage statistique, et offre une description très simplifiée de leur fondement mathématique. Cependant, on peut en tirer plusieurs leçons sur la nature des données et de leur traitement.

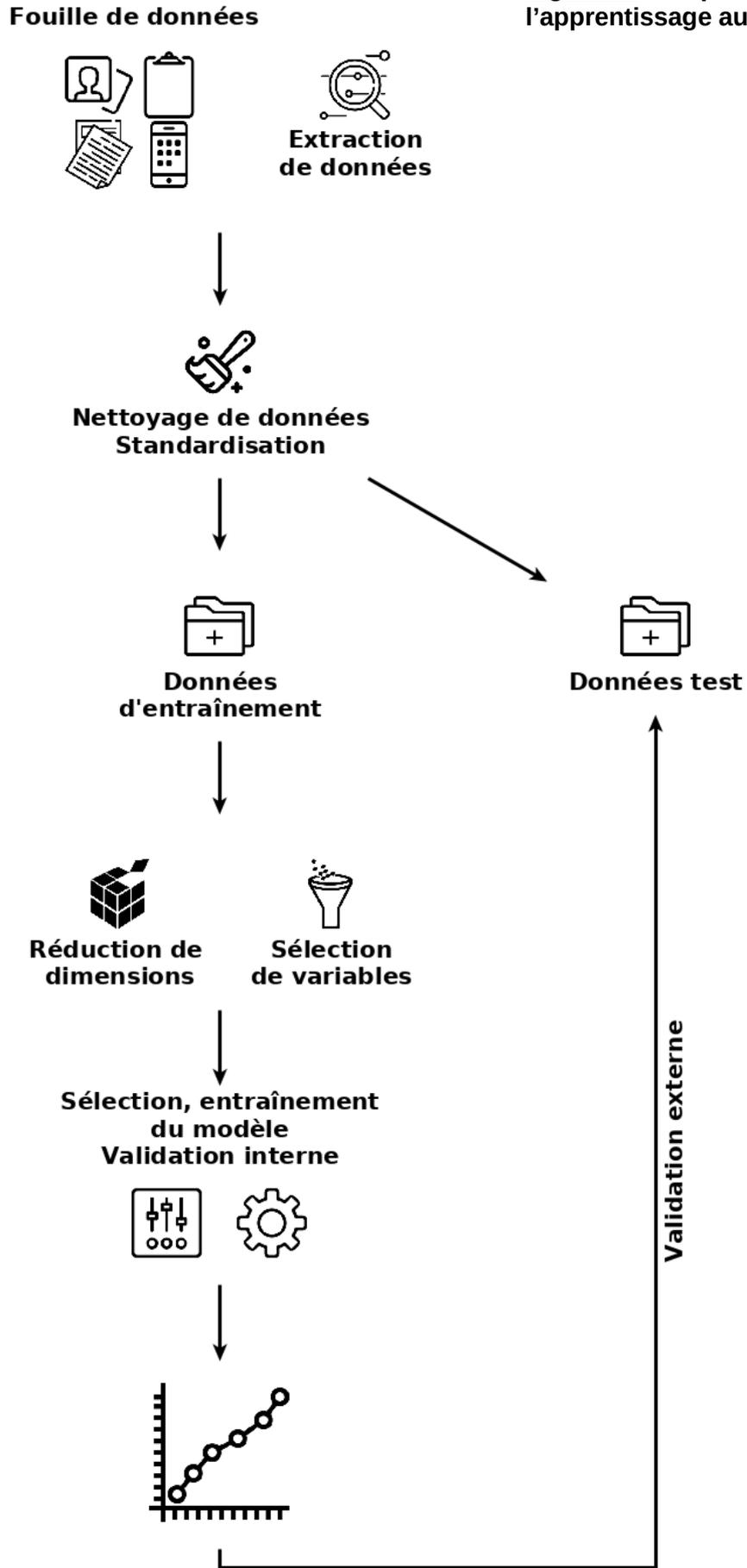
- Premièrement, on ne peut juger la signification d'un résultat que si on a une compréhension du fonctionnement, et surtout des limites, de la technique qui l'a produit. Ce point mérite d'être clairement énoncé, même s'il peut paraître évident.
- Le leitmotiv de cette discussion des techniques d'apprentissage est le compromis biais-variance, et la manière dont on cherche un équilibre entre les deux, entre complexité et intelligibilité.
- Ce balancement nous ramène au dilemme de base évoqué en introduction : la tension à négocier entre généralisation et personnalisation. Le postulat derrière les initiatives de Big Data est que l'accumulation de davantage de données permettra de surmonter cette tension, de la faire disparaître. Mais finalement, on ne fait que déplacer ce dilemme sur un autre plan, statistique et mathématique. Pour citer les auteurs de *An Introduction to Statistical Learning* :

« Thus, we see that new technologies that allow for the collection of measurements for thousands or millions of features are a double-edged sword: they can lead to improved predictive models if these features are in fact relevant to the problem at hand, but will lead to worse results if the features are not relevant. *Even if they are relevant, the variance incurred in fitting their coefficients may outweigh the reduction in bias that they bring.* » (James et al., 2013, italiques rajoutées)

Il ne faut pas se leurrer sur la portée des outils d'apprentissage statistique, aussi puissants et efficaces qu'ils puissent se montrer dans certains domaines. Ils n'offrent pas de certitudes face à la complexité, mais des méthodes pour la gérer. Ils ne dévoilent pas la réalité objective, mais des interprétations plus ou moins contextualisées. Dans le meilleur des cas, le *Big Data* et ces outils nous permettent peut-être de mieux cerner ce que *nous ne savons pas* et où focaliser nos efforts de recherche, en montrant par exemple, les failles d'un modèle sur un test de validation ou les variations de performance d'un modèle dans le monde réel – seulement si nous sommes assez avertis pour y prêter attention.

Figure ci-après réalisée avec l'aide d'icônes sous licence Creative Commons (Freepik, s. d.; phatplus, s. d.)

Figure 13 : Étapes clés de l'apprentissage automatique



c. Initiatives de partage des données

L'une des caractéristiques du Big Data abordées précédemment (chapitre I.1.) est l'existence de larges bases de données interconnectées, dont l'intérêt est d'être interrogées à plusieurs reprises. D'ailleurs cette utilisation est implicite dans la manière dont de nombreuses bases, surtout commerciales, sont constituées, brassant par défaut des masses de données, plutôt que de spécifier et limiter les informations collectées à un objectif précis. Même les données collectées pour répondre à une question précise, comme un essai clinique ou une étude expérimentale, peuvent potentiellement receler des informations utiles à d'autres problématiques. À une ère d'optimisation, judicieuse ou exagérée, des ressources publiques, l'élargissement de l'accès à ces données, et la compilation de grandes bases à partir de données déjà existantes, deviennent une priorité. Les notions d'Open Data (données ouvertes) et *open access* (accès libre) répondent à ces impératifs, et sont adoptées par de nombreuses institutions publiques.

On distingue les données agrégées, au niveau d'un territoire par exemple, et les données individualisées et détaillées, qu'elles appartiennent à des personnes humaines, ou proviennent d'études animales ou cellulaires. L'idée du partage est basée sur les principes suivants, résumés dans l'acronyme FAIR (« About BD2K | Data Science at NIH », s. d.; Hagstrom, 2014) : *Findable* (trouvable, interrogeable), *Accessible*, *Interoperable* (intercompatible et compréhensible), and *Reusable* (réutilisable). Les données peuvent être mises à disposition par les organismes publics qui les génèrent, comme sur data.gouv.fr ou le SNDS, ou réclamées par un organisme public comme condition au financement ou à la validation d'un projet. C'est le cas des projets de recherche soutenus par le programme Horizon 2020 de la Commission Européenne, et par le NIH (Iatropoulou, s. d.; « NIH Data Sharing Information », s. d.). Les deux institutions exigent le partage des résultats de projets de recherche par la mise à disposition gratuite des publications en lien avec le projet. Elles proposent également un partage des données « par défaut », incitant les groupes de recherche à télécharger, autant que possible, leurs données brutes sur des plateformes de partage (Zenodo ou OpenAIRE pour l'UE), et à justifier les limites au partage. Dans le même esprit, l'enregistrement des essais cliniques, dans des bases de données librement accessibles et incluant souvent un résumé des résultats, est largement devenu obligatoire (« ClinicalTrials.gov », s. d.; « EU Clinical Trials Register », s. d.; « WHO | ICTRP

International Clinical Trials Registry Platform », s. d.). L'accès gratuit aux données peut être totalement ouvert, avec des fichiers en téléchargement libre direct, ou soumis à un accord préalable, sous la forme de demandes ponctuelles ou d'un accès restreint à certaines catégories de personnes approuvées. Ce format permet d'avoir plusieurs niveaux de sécurité, de protection de la vie privée et de la propriété intellectuelle, en fonction de la sensibilité des données. Les SNDS fonctionne selon un tel système de gradation des données, ainsi que la SAIL Databank au Royaume-Uni. En recherche, il existe de nombreuses initiatives de mise en commun ou de partage des données, à travers les consortiums de recherche, ou des banques de données, comme celles recensées par Woo et al. en neuroimagerie (Tableau 2, Woo, Chang, Lindquist, & Wager, 2017).

En mettant à disposition des données très variées à de nombreux acteurs, ces initiatives facilitent, dans le meilleur des cas, la dissémination du savoir, la vérification et la validation de travaux scientifiques, la responsabilisation et la transparence du service public, et un engagement citoyen plus éclairé. Mais l'Open Data soulève aussi des questions légales et éthiques de propriété intellectuelle, d'usurpation de données ouvertes à des fins commerciales, d'atteinte à la vie privée dans le cas de données personnelles identifiables. Au vu de cette ambivalence, les cadres réglementaires qui s'adaptent à la nature des données et aux implications éthiques, ou qui offrent différents degrés d'application, sont peut-être plus sensés que des règles strictes et générales, comme le montre l'échec de l'imposition du partage de données proposé dans le réseau NHS England (Boseley, 2016).

Resource	Initials	Clinical groups	Imaging modality	Web address
Data-sharing platforms				
International Data-sharing Initiative	INDI			http://fcon_1000.projects.nitrc.org/
Neuroimaging Informatics Tools and Resources Clearinghouse	NITRC			http://www.nitrc.org/
Laboratory of Neuro Imaging Image & Data Archive	LONI IDA			https://ida.loni.usc.edu/login.jsp
Collaborative Informatics and Neuroimaging Suite	COINS			http://coins.mrn.org/
National Alliance for Medical Image Computing	NA-MIC			http://wiki.na-mic.org/Wiki/index.php/Main_Page
Open functional Magnetic Resonance Imaging	OpenfMRI			https://openfmri.org/
Consortium or repository				
Autism Brain Imaging Data Exchange	ABIDE	Autism	rs-fMRI	http://fcon_1000.projects.nitrc.org/indi/abide/
Attention Deficit Hyperactivity Disorder-200	ADHD-200	ADHD	rs-fMRI	http://fcon_1000.projects.nitrc.org/indi/adhd200/
Alzheimer's Disease Neuroimaging Initiative	ADNI	Alzheimer's	Multimodal, including sMRI, fMRI, PET	http://adni.loni.usc.edu/ ; accessible through IDA
Australian Imaging, Biomarkers & Lifestyle Flagship Study of Aging	AIBL	Alzheimer's	Multimodal, including sMRI, PET	https://aibl.csiro.au/ ; accessible through IDA
Open Access Series of Imaging Studies	OASIS	Alzheimer's	sMRI	http://oasis-brains.org/
International Study to Predict Optimized Treatment for Depression	iSPOT-D	Depression	Multimodal, including EEG, dMRI, fMRI, sMRI	http://www.brainresource.com/home.html
MIND Clinical Imaging Consortium	MCIC	Schizophrenia	Multimodal, including dMRI, fMRI, sMRI	http://coins.mrn.org
Center of Biomedical Research Excellence	COBRE	Schizophrenia	sMRI, rs-fMRI	http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html
function Biomedical Informatics Research Network	fBIRN	Schizophrenia	fMRI, sMRI	http://www.schizconnect.org/
Parkinson's Progression Markers Initiative	PPMI	Parkinson's	Multimodal, including dMRI, fMRI, sMRI, PET	http://www.ppmi-info.org/
Pain and Interoception Imaging Network	PAIN	Pain	Multimodal, including dMRI, fMRI, sMRI	http://www.painrepository.org/
OpenPain	Openpain	Pain	fMRI	http://www.openpain.org

Note: dMRI, fMRI and sMRI refer to diffusion-weighted, functional and structural MRI, respectively; rs-fMRI refers to resting-state fMRI.

Tableau 2 : Initiatives de partage de données d'imagerie et consortiums de recherche, pour les troubles neurologiques et psychiatriques
(Woo, Chang, Lindquist, & Wager, 2017)

II – Utilisations du Big Data en psychiatrie⁷

Ce chapitre examine, à travers une sélection de publications, la place du Big Data en psychiatrie. Mais le choix de publications provenant de revues biomédicales à comité de lecture ne donne qu'une vision très partielle de cette question. En effet, le premier chapitre mentionnait la multiplicité des sources de référence et de communications techniques dans le domaine des sciences des données. En raison de cette limitation du choix de sources, imposé par souci de synthèse, d'intelligibilité et de vérifiabilité, ce chapitre doit être conçu non comme un exposé exhaustif de l'utilisation du Big Data en psychiatrie, mais plutôt comme un aperçu de travaux notables apparaissant sous un format reconnu par la communauté médicale.

La difficulté d'une synthèse exhaustive des travaux de Big Data en psychiatrie est exacerbée par la définition floue du Big Data, et l'hétérogénéité du vocabulaire associé au concept. *Big Data, machine learning, data-driven, deep learning, data mining, clustering, prediction* etc, sont certains termes employés dans les titres ou mots-clés de publications pour signifier une approche Big Data telle que nous l'avons délimitée au premier chapitre. Pour cette raison, certaines revues de la littérature choisissent plutôt de se focaliser sur un aspect particulier : une technique comme l'apprentissage profond dans le revue collaborative de Ching et al. (Ching et al., 2018), un objectif comme la prédiction de l'efficacité des traitements en psychiatrie (Gillan & Whelan, 2017), ou une pathologie comme la dépression (Patel, Khalaf, & Aizenstein, 2016) ou le trouble bipolaire (Librenza-Garcia et al., 2017).

Néanmoins, quatre grands thèmes se dessinent au sein des travaux de Big Data en psychiatrie, qui sont en parallèle avec les étapes de la pratique clinique, et qui peuvent, bien entendu, s'entrecouper et se mélanger :

- découverte : chercher de nouvelles façons de comprendre ou expliquer la maladie, les mécanismes physiopathologiques et psychopathologiques,

⁷ Tout au long de ce chapitre, le terme de « précision » sera utilisé pour signifier l'*accuracy* telle qu'elle est communément rapportée dans ces études, comme mesure d'efficacité des modèles de classification et de prédiction, en sachant que : $accuracy = (\text{vrais positifs} + \text{vrais négatifs}) / (\text{cas} + \text{témoins})$ c'est-à-dire la proportion de prédictions/classifications correctes, qu'elles soient négatives ou positives.

- classification : trouver des critères plus pertinents pour désigner le trouble dont le patient est atteint et grouper les patients selon leurs troubles, c'est-à-dire dépister et diagnostiquer,
- prédiction : pouvoir déterminer, pour un patient donné, l'évolution naturelle d'un trouble ou la réponse à un traitement, donc faire un choix thérapeutique et connaître le pronostic,
- suivi : développer des manières différentes de poursuivre le travail clinique, qu'il soit diagnostique ou thérapeutique, avec le patient.

Pour identifier les articles d'intérêt, une recherche a été effectuée sur PubMed et Google Scholar avec les mots-clés « Big Data », « data science », « machine learning » ou « deep learning », associés aux mots-clés « psychiatry » ou « mental health ». Des articles supplémentaires étaient identifiés après lecture de revues de la littérature et un parcours de leurs références. Ceci n'étant pas une revue systématique, les articles retrouvés n'étaient pas lus dans leur totalité. En effet, des articles pertinents étaient repérés en parcourant les titres retrouvés par les moteurs de recherche, et après lecture de leurs abstracts, les articles étaient sélectionnés pour discussion en fonction de leur valeur illustrative des différentes facettes et utilisations du Big Data dans une variété de contextes et de problématiques psychiatriques, avec une attention particulière à la recherche translationnelle et aux applications cliniques pratiques.

1. Découvrir

Les critiques récurrentes, dans la littérature médicale, envers l'état actuel du savoir en psychiatrie, concernent la subjectivité des concepts et des outils d'exploration, et le manque de validité neurobiologique du cadre nosologique (Insel et al., 2010; Kapur, Phillips, & Insel, 2012). Selon les auteurs, ces défaillances entravent une meilleure intégration entre les neurosciences et la clinique, et par conséquent les avancées que cette synthèse apporterait.

Les principaux conduits d'observation en psychiatrie sont le comportement et le discours, qui peuvent donner accès, de manière plus ou moins indirecte et oblique, aux pensées et aux émotions, et à leurs perturbations qui constituent les symptômes. L'approche de certains auteurs est de rendre ces observations plus « objectives ». Ben-Zeev et al. proposent CrossCheck, un programme de suivi pour patients psychotiques

(Ben-Zeev et al., 2017), basé sur une application smartphone qui interroge régulièrement les utilisateurs sur leur état de santé, et enregistre en continu des informations sur leur comportement par l'intermédiaire de plusieurs capteurs : accéléromètre pour le niveau d'activité physique, localisation géographique pour connaître leurs déplacements, microphone intégré pour évaluer la durée et la fréquence de conversations ambiantes etc. Ces mesures automatiques donneraient, selon les auteurs, accès à une vision plus fine et plus quantifiable du comportement des patients dans leur propre milieu et à toute heure de la journée, sans le biais et la médiation d'un interrogatoire clinique, augmentant notre compréhension de la pathologie et de son évolution. Un essai clinique de cette application est en cours, et les auteurs publient ici cinq cas de patients suivis pendant un an. Ils décrivent, entre les périodes d'hospitalisation, une grande variabilité de comportements parmi les cinq patients, certains signalant la dégradation de leur état dans un questionnaire pluri-hebdomadaire, d'autres présageant une décompensation par des modifications du comportement enregistré (changement de localisation géographique, augmentation ou au contraire diminution de parole, arrêt d'utilisation de l'application).

Dans le même esprit, d'autres auteurs proposent d'aborder le langage et le discours des patients d'une manière différente pour comprendre la pathologie psychiatrique. Dans une revue sélective de la littérature, Elvevåg et al. plaident pour une conceptualisation plus sophistiquée du langage dans les domaines RDoC, et dans la recherche et l'évaluation en psychiatrie (Elvevåg et al., 2016), citant la nécessité d'une approche pluridisciplinaire et de modèles expérimentaux s'appuyant sur les avancées en psycholinguistique, en apprentissage automatique et en traitement automatique du langage, ainsi que sur l'analyse de nouvelles sources de données (acoustique et prosodie d'une voix enregistrée par exemple). Selon ces auteurs, l'analyse objective lexicale, sémantique, syntaxique et phonétique du discours des patients, en combinaison avec d'autres marqueurs comme l'imagerie cérébrale ou le séquençage génétique, mènerait à la découverte de processus psychopathologiques ou étiologiques en psychiatrie.

Plusieurs études mettent en œuvre cette idée. Mota et al. pratiquent une analyse quantitative du discours en représentant le discours comme un réseau de mots, puis en calculant les différentes propriétés et la structure de ce réseau (nombre de nœuds et de liens, diamètre du réseau, nombre de sous-réseaux, etc, Figure 14). Ils comparent ainsi, sur la base de ces mesures, les discours de patients (schizophrénie, trouble bipolaire) et de personnes sans trouble psychiatrique identifié. Leurs deux études, utilisant des transcriptions du discours de personnes racontant leur dernier rêve (Natália B. Mota,

sémantique latente, ou LSA, qui se base sur la fréquence d'apparition de termes dans différents documents pour relier les termes et les documents à des concepts, et estimer en quelque sorte le signifié. Ils identifient la « cohérence sémantique » (similitude sémantique) entre phrases consécutives, et le nombre de pronoms relatifs, comme facteurs d'intérêt, interprétant la cohérence comme un reflet du niveau de « désordre » dans un texte, et le nombre de pronoms comme un reflet de la richesse/pauvreté du discours. Le fil conducteur entre ces études est la quantification automatisée des caractéristiques du langage, possiblement dans le but de distinguer des motifs ou des schémas psychopathologiques plus fins que les symptômes actuels, mais aussi peut-être dans l'espoir de rendre l'évaluation diagnostique moins subjective, moins dépendante de l'observateur et plus automatisée.

Identifiant une source de langage autre que le discours recueilli lors d'interrogatoires cliniques, Guntuku et al. passent en revue les études utilisant les données de réseaux sociaux pour détecter les troubles dépressifs et d'autres troubles psychiatriques (Guntuku, Yaden, Kern, Ungar, & Eichstaedt, 2017). Ils recensent des analyses portant à la fois sur les caractéristiques sémantiques de tweets ou de billets publiés dans des forums ou des pages Facebook, et sur les méta-données enregistrant les horaires et le nombre de connexions aux réseaux, les interactions avec d'autres utilisateurs etc. Ils distinguent les études utilisant des données disponibles publiquement, des études recrutant des utilisateurs de réseaux sociaux et sollicitant un accès à leurs données privées. Les études se différencient aussi par la manière dont le diagnostic présumé est établi : par l'administration de questionnaires diagnostiques en ligne, par la recherche d'un diagnostic déclaré par l'utilisateur, par l'appartenance à un groupe en ligne désigné, ou par l'extrapolation d'un diagnostic à partir des différentes données. Les auteurs ne recensent aucune étude où le diagnostic est établi par un interrogatoire et un examen clinique. La taille des échantillons dans ces études varie d'une centaine à plus de 20 000, et les résultats sont inégaux. Une étude identifie, par exemple, des changements de niveaux d'activité et de langage (émotions négative, termes liés à la dépression, retour sur soi, moins d'interactions) sur Twitter chez les utilisateurs déprimés dans l'année précédant un épisode dépressif (De Choudhury, Gamon, Counts, & Horvitz, 2013). Une autre étude, basée sur des forums reddit, distingue certaines caractéristiques d'utilisateurs qui présenteront des idées suicidaires : un langage appauvri, la perte d'espoir, l'anxiété, un sentiment de solitude et d'impulsivité, un repli sur soi et une diminution des engagements sociaux (De Choudhury, Kiciman, Dredze, Coppersmith, & Kumar, 2016).

Mais on trouve aussi un exemple illustrant la difficulté d'interpréter les résultats issus de ces données non-contrôlées et multi-dimensionnelles. Dans une analyse de tweets publics, une étude retrouve une grande superposition entre le langage caractérisant les utilisateurs déclarant un trouble psychiatrique (dépression, syndrome post-traumatique) et le langage prédisant le type de personnalité et l'âge, signalant peut-être un facteur confondant entre le type de personnalité, l'âge et la propension à révéler un diagnostic sur Twitter, plutôt qu'un lien psychopathologique ou étiologique (Preoțiu-Pietro et al., 2015). On voit donc que ces différentes études exploitent, avec plus ou moins de succès, les données des réseaux sociaux à la fois comme marqueurs du comportement (méta-données) et comme marqueurs de la pensée et des émotions (analyse sémantique), ouvrant potentiellement la voie vers une meilleure caractérisation psychopathologique et étiologique, et vers des outils de prédiction et de classification, comme nous verrons plus loin (sous-chapitres II.2 et II.3). Guntuku et al. mentionnent également l'intérêt potentiel de ces données comme source d'informations biographiques pour l'analyse des circonstances de vie associées aux troubles psychiatriques.

Au-delà de l'expression verbale comme fenêtre sur le psychisme, d'autres auteurs examinent le langage par la manière dont certains concepts sont représentés chez les personnes avec et sans trouble psychiatrique, grâce à l'analyse neurosémantique. Just et al. utilisent un schéma expérimental d'IRM fonctionnelle (IRMf) avec un stimulus sémantique, pour rechercher des altérations de représentations conceptuelles chez des personnes présentant des idées suicidaires (Just et al., 2017) et avec autisme (Just, Cherkassky, Buchweitz, Keller, & Mitchell, 2014). Les auteurs se basent sur des travaux en neurosciences cognitives et en apprentissage automatique qui retrouvent des schémas d'activation en IRMf caractéristiques de certains concepts et peuvent, dans une certaine mesure, prédire ces schémas pour de nouveaux concepts : lorsque des sujets sont chargés de focaliser leur pensée sur un concept désigné par un mot qui leur est présenté, les schémas d'activation en IRMf qui résultent de cette activité mentale, constituent une sorte de « signature neuronale » du concept, qui semble identifiable, reproductible, prédictible et cognitivement cohérente chez les sujets sains (Mitchell et al., 2008). Dans leurs deux études, Just et al. testent les représentations de trente mots liés au suicide et aux émotions négatives et positives (*suicide, overdose, desperate, carefree, good, cruelty, guilty, trouble...*) chez 17 patients suicidaires et 17 sujets sains, et de huit mots liés à l'interaction sociale (*compliment, insult, hate, hug, encourage...*) chez 17 sujets autistes et 17 sujets sains. Ils identifient les concepts et les localisations cérébrales pertinentes pour

distinguer les patients, par des techniques de réduction/sélection de dimension (voir I.3.b. p. 45-46, dans ces études : *factor analysis*, et une procédure similaire à *stepwise regression*). Dans l'étude de patients suicidaires, les termes *death*, *cruelty*, *trouble*, *carefree*, *good*, *praise* apparaissent les plus pertinents, ainsi que l'activation dans les régions médiale frontale gauche, cingulaire antérieure et temporale droite (sujets sains), et inférieure frontale et pariétale gauches (sujets suicidaires). L'étude de patients autistes montre une absence d'activation du gyrus cingulaire postérieur et du précuneus, par rapport aux sujets sains, en réponse aux concepts relationnels. Ici le langage est utilisé comme paradigme expérimental pour accéder à un fonctionnement cognitif aberrant, montrant comment les techniques d'apprentissage automatique sont utilisées pour à la fois découvrir des schémas généraux et prédire des différences individuelles, et pour gérer des données à haute-dimension en réduisant le « bruit » pour trouver le « signal ». Ces différentes approches au langage constituent des niveaux croissants d'abstraction, au sens de conceptualisation et de réduction, mais on peut se demander si elles ne mènent pas à une autre sorte d'abstraction, celle de l'exclusion du discours et du sujet⁸.

D'autres études développent le volet cognitif en examinant le fonctionnement cognitif au sens large ou en relation à des symptômes cardinaux, dans différents troubles psychiatriques. Wu et al. cherchent à mieux caractériser le trouble bipolaire, en examinant le fonctionnement cognitif couplé à une visualisation de la substance blanche par IRM de diffusion (Wu et al., 2017). Avec l'aide d'un algorithme d'apprentissage non-supervisé (*k-means clustering*), ils analysent les résultats de tests neurocognitifs (Cambridge Neurocognitive Test Automated Battery ou CANTAB) administrés à 70 patients bipolaires, et distinguent deux sous-groupes de patients. Cherchant un substrat neurobiologique à cette différence phénotypique, ils construisent un modèle d'apprentissage supervisé (*Elastic Net* avec modèle linéaire) à partir des IRM de diffusion des patients (utilisant les valeurs d'anisotropie fractionnelle et de diffusivité moyenne comme variables), évaluant la capacité du modèle et de ces variables à classifier les deux groupes, et les patients par rapport à des sujets sains. Ainsi, ils identifient plusieurs tâches d'intérêt dans la batterie neurocognitive, liées au contrôle affectif et cognitif, à la prise de décision, à la mémoire de travail et à la planification, ainsi qu'une région d'intérêt en imagerie, les faisceaux de substance blanche frontale. Ces travail peut se résumer comme une analyse multi-modale, par apprentissage supervisé et non-supervisé, du fonctionnement cognitif pour mieux caractériser la variabilité phénotypique au sein d'un trouble psychiatrique.

8 Cette idée sera développée au chapitre III de cette thèse.

Toujours dans l'exploration cognitive, mais avec deux approches différentes, Wensing et al. et Brodersen et al. recherchent le corrélât cognitif et neurobiologique de groupements de symptômes schizophréniques : la dissociation psychique (Wensing et al., 2017) et les symptômes négatifs (Brodersen et al., 2014). Les premiers synthétisent les résultats d'une méta-analyse de 18 études d'imagerie fonctionnelle (tâche-dépendante ou de repos), à l'aide d'une analyse par *Activation Likelihood Estimation* qui identifie les régions convergentes dans l'imagerie entre différentes études. Le résultat le plus robuste qu'ils retrouvent en lien avec la dissociation psychique, correspond aux modifications d'activité dans le gyrus temporal gauche (moyen postérieur et supérieur). Ils tentent ensuite d'inférer les processus cognitifs en jeu, en imputant les réseaux fonctionnels cérébraux liés à ces régions, et l'activité cognitive de ces réseaux, grâce à l'exploitation de la base BrainMap (« brainmap.org | Home », s. d.), une base compilant des études publiées en neuro-imagerie fonctionnelle et morphologique, avec les résultats référencés par coordonnées standardisées (espace Talairach ou MNI) et des outils de partage et de méta-analyse. Wensing et al. proposent ainsi un mécanisme cognitif sous-tendant la dissociation psychique – l'association de processus cognitifs liés au langage, à la parole, à la perception auditive et à la cognition sociale – soulignant l'implication non seulement du langage, mais aussi de fonctions cognitives supérieures. Cette étude illustre les aspects de partage des données et du code d'algorithmes, la compilation de grandes bases référencées et interrogeables, et l'utilisation d'algorithmes pour imputer de nouvelles associations entre des données initialement non liées.

Brodersen et al. utilisent une tâche de mémoire de travail en IRM fonctionnelle pour analyser le fonctionnement cognitif de 42 patients schizophrènes et 42 sujets sans trouble (Brodersen et al., 2014). Essayant d'aller au-delà de la connectivité anatomique et fonctionnelle, ils modélisent pour cette tâche, un réseau visuel-pariétal-préfrontal, grâce à la technique de *Dynamic Causal Modeling* (DCM), une méthode de modélisation probabiliste (Bayésienne) pour inférer la connectivité effective d'un réseau, c'est-à-dire inférer des liens de cause et effet entre l'activation de différentes régions, et non seulement une corrélation temporelle ou anatomique (Friston, Harrison, & Penny, 2003; Heinzle & Stephan, 2018). Les paramètres issus de ce modèle sont ensuite intégrés (*generative embedding*) à des modèles d'apprentissage supervisé et non supervisé pour classifier les patients par rapport aux sujets sains, et pour découvrir de nouveaux sous-groupes de patients (*patient stratification*). L'analyse non-supervisée des patients seuls trouve trois clusters de patients avec une connectivité du réseau « mémoire de travail »

(visuel-pariétal-préfrontal) différente. Ils interprètent ces différences comme indicatives d'une intégration fonctionnelle plus resserrée entre les trois régions du réseau chez les patients du premier et troisième cluster, et d'un découplage du cortex pariétal et préfrontal chez les patients du deuxième cluster. Recherchant un corrélat clinique à cette différence d'architecture, les auteurs observent que les trois sous-groupes se distinguent aussi par leur scores à l'échelle de syndrome négatif de la PANSS (Positive and Negative Syndrome Scale) : le premier sous-groupe ayant un score inférieur à la moyenne du groupe de patients, le deuxième cluster ayant un score proche de la moyenne, et le troisième se situant au-dessus de la moyenne. Comme Wu et al. avec le trouble bipolaire, ces auteurs utilisent l'apprentissage automatique à plusieurs niveaux, pour explorer les dysfonctionnements cognitifs dans la schizophrénie, et vont plus loin en essayant de les caractériser cliniquement (PANSS) et leur imputer un mécanisme.

Ces descriptions cognitives signent l'insatisfaction actuelle avec les nosographies dominantes et la recherche d'un autre cadre nosologique, comme le projet RDoC cité en introduction. L'étude de Gillan et al. va plus explicitement dans ce sens, en partant d'un processus cognitif spécifique et en essayant de trouver la dimension de symptômes qui y est associée, plutôt que de partir d'un trouble psychiatrique donné et d'en rechercher les altérations cognitives (Gillan, Kosinski, Whelan, Phelps, & Daw, 2016). Faisant l'hypothèse que les compulsions, observables dans les troubles obsessionnels-compulsifs (TOC) et les addictions, sont dues à un déficit en *goal-directed control* (contrôle ou apprentissage orienté vers un but), induisant un manque de flexibilité cognitive, ces auteurs notent que ce déficit a aussi été décrit dans la schizophrénie et l'anxiété sociale. Selon eux, cette non-spécificité ne remet pas en question le lien psychopathologique entre ce processus cognitif et le symptôme de compulsion, mais reflète plutôt le manque de validité des catégories diagnostiques. Leur étude se base donc sur l'administration en ligne d'une tâche cognitive estimant le *goal-directed control*, et de plusieurs questionnaires cliniques estimant des dimensions de symptômes, à deux groupes de personnes issues de la population générale, recrutées sur la plateforme de micro-travail et de *crowdsourcing*, Amazon Mechanical Turk (548 sujets dans le premier groupe, 1413 sujets dans le deuxième groupe). Les résultats sont analysés par des méthodes d'apprentissage statistique (régression logistique supervisée avec *Elastic Net regularization*, et analyse factorielle non-supervisée). Les auteurs trouvent, dans le premier groupe, une association significative entre la sévérité des scores sur le questionnaire TOC et un déficit en *goal-directed control*, association qu'ils ne retrouvent pas avec les scores de dépression ou

d'anxiété. Ce résultat est validé dans le deuxième groupe indépendant, et l'association se retrouve également avec les questionnaires évaluant les troubles du comportement alimentaire, l'impulsivité et l'alcool-dépendance. Par une analyse non-supervisée des différents items de chaque questionnaire, les auteurs identifient trois dimensions de symptômes évalués par ces questionnaires, qu'ils nomment « comportement compulsif/pensées intrusives », « retrait social », « anxiété-dépression ». La dimension significativement associée au déficit en *goal-directed control* est la dimension « comportement compulsif/pensées intrusives », qui contient des items sur les compulsions et les obsessions provenant des questionnaires TOC, addiction et trouble du comportement alimentaire. En tirant parti d'un large échantillon de sujets, de résultats multidimensionnels, d'une modélisation par apprentissage statistique, Gillan et al. infèrent un lien spécifique entre une dimension de symptômes et un processus cognitif particulier au sein d'une population générale, majoritairement infra-clinique, proposant une manière différente de classer les symptômes et syndromes, et d'en identifier un mécanisme.

La recherche d'un mécanisme, qui se confond avec la poursuite d'une explication étiologique, jusque-là insaisissable en psychiatrie, peut prendre d'autres chemins. On en voit des exemples dans les travaux de Hazlett et al. et de Xiong et al. qui examinent les troubles autistiques par le biais du développement et de la biologie moléculaire (Hazlett et al., 2017; Xiong et al., 2015). À partir d'une large cohorte d'enfants suivis pour une étude prospective multicentrique (Infant Brain Imaging Study, IBIS Network), Hazlett et al. examinent les différences entre les IRM morphologiques de 106 enfants à haut-risque familial d'autisme et 42 enfants à bas-risque, à l'âge de six, douze et 24 mois. Parmi les enfants à haut-risque, quinze enfants reçoivent un diagnostic de trouble autistique à 24 mois, tous les autres enfants (91 à haut-risque et 42 à bas-risque) ne présentant pas de trouble à cet âge. Les auteurs modélisent les trajectoires de croissance cérébrale (volume cérébral total, surface et épaisseur corticale) en fonction de l'âge, pour chaque groupe d'enfants, grâce à une fonction linéaire en morceaux (*piecewise longitudinal mixed model*). Par rapport aux enfants indemnes, ils observent une expansion plus rapide de la surface cérébrale dans la première année, et du volume cérébral total dans la deuxième année, chez les enfants diagnostiqués autistes à 24 mois, suggérant un processus développemental distinct chez ces enfants. Le taux de croissance du volume total dans la première année est également corrélé à la sévérité de certains scores cliniques d'autisme chez ces enfants à deux ans. Les auteurs utilisent ensuite un algorithme entraîné par apprentissage profond sur 312 mesures (sexe, volume intra-crânial, surface et épaisseur

de 39 régions corticales bilatérales) dans un groupe plus élargi de 179 enfants à haut-risque, et distinguent ceux qui ont un diagnostic d'autisme avec une spécificité de 95 % et une sensibilité de 88 %, notant que la mesure qui semble contribuer le plus à la décision algorithmique est la surface cérébrale à six mois. Cette étude est un bon exemple de la haute-dimensionalité, où le nombre de variables/mesures est supérieur au nombre de sujets, et du déploiement de l'apprentissage automatique, à la fois pour l'inférence où on s'intéresse à la forme et l'intelligibilité de la fonction (modélisation des trajectoires de croissance), et pour la prédiction où on recherche la précision du résultat (détection des enfants autistes).

L'approche développementale morphologique de Hazlett et al. peut être enrichie par une approche de connectivité fonctionnelle, telle qu'elle est mise en œuvre par Dosenbach et al. dans leur étude de la maturité cérébrale (Dosenbach et al., 2010). En analysant les données d'IRM fonctionnelle de repos, auprès de 238 individus sains âgés de sept à trente ans, les auteurs développent un index mesurant la maturation de la connectivité fonctionnelle cérébrale. Cet index montre une bonne précision (supérieure à 90%) dans la classification par apprentissage automatique d'IRM individuelles en deux groupes, « enfant » et « adulte », lorsqu'il est testé et validé sur deux échantillons indépendants de l'échantillon initial. Examinant l'évolution de la connectivité fonctionnelle de l'enfance à l'âge adulte, les auteurs observent que la maturation cérébrale semble marquée par un renforcement de la connectivité entre régions distantes, et l'affaiblissement de la connectivité locale, correspondant probablement à une meilleure intégration de réseaux fonctionnels impliquant des régions cérébrales distinctes. Ils soulignent l'intérêt de cette approche dans l'exploration de troubles neurologiques ou psychiatriques. À la différence de l'étude de Hazlett et al., il ne s'agit pas ici d'un suivi longitudinal, mais de l'imputation d'un processus développemental commun à partir de données provenant de différents individus d'âges distincts.

Se démarquant d'autres techniques d'étude génétique qui reposent sur une annotation fonctionnelle du génome ou sur une large population pour détecter des polymorphismes d'intérêt, Xiong et al. proposent un algorithme développé par apprentissage profond (*Bayesian deep learning algorithm*), nommé SPANR (splicing-based analysis of variants), qui modélise la régulation de l'épissage (Xiong et al., 2015). Alimenté par des bases génomiques et transcriptionnelles de référence, cet algorithme « apprend » l'effet de différentes séquences d'ADN nucléaire sur la transcription et l'épissage, dans une variété de contextes génomiques et cellulaires normaux. Ensuite,

lorsqu'il est appliqué à toute séquence d'ADN humain contenant trois exons, il prédit pour un type cellulaire donné, le pourcentage d'ARN messagers épissés avec l'exon central (Figure 15). Ce modèle est validé par les auteurs sur plusieurs essais (avec un succès variable selon l'essai), puis utilisé pour prédire l'effet de polymorphismes associés à certaines pathologies. Les auteurs décrivent, par exemple, l'utilisation de SPANR pour analyser des génomes entiers séquencés du tissu cérébral de cinq personnes autistes et douze sujets sains. Parmi les polymorphismes repérés sur ces génomes, SPANR identifie ceux qui causent potentiellement des anomalies d'épissage : 171 chez les sujets autistes, 249 chez les sujets sains. Chez les personnes autistes, une plus grande proportion de ces polymorphismes correspond à des gènes exprimés dans le tissu cérébral (46 gènes ou 27 %, par rapport à 33 gènes ou 13 % chez les sujets sains), identifiant deux gènes précédemment impliqués dans l'autisme, et d'autres gènes associés au développement ou au fonctionnement du système nerveux. SPANR offre davantage de possibilités que la simple association statistique entre un allèle et une pathologie : cet outil ouvre une porte sur l'étude de l'impact fonctionnel d'un polymorphisme, et sur la découverte de nouvelles associations rares ou spontanées, difficiles à détecter par les autres méthodes d'exploration génétique. La construction de SPANR repose sur et alimente plusieurs aspects du Big Data, notamment la disponibilité de larges bases génétiques de référence en accès ouvert, l'utilisation de l'apprentissage profond pour trouver des schémas/motifs répétés dans ces données et pour faire des prédictions, et le partage ouvert de l'outil développé.

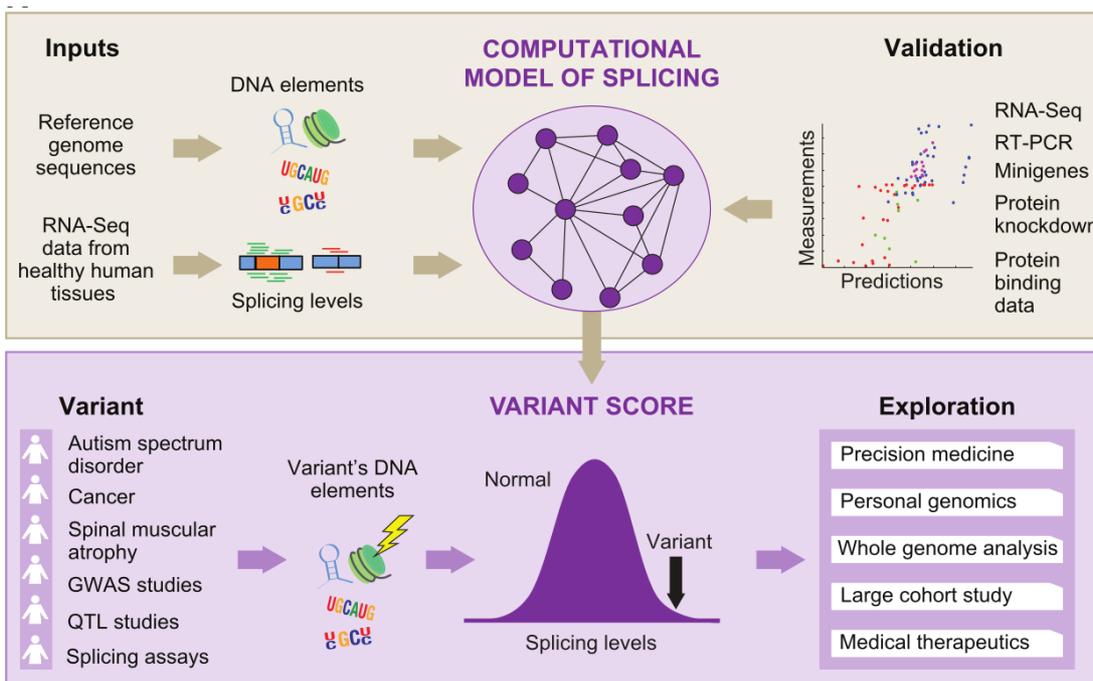


Figure 15 :
Protocole
d'apprentissage
de l'outil
SPANR
 (Xiong et al.,
 2015)

La majorité des études présentées ici ont des échantillons de taille modeste ou moyenne, loin des milliers d'individus auxquels l'appellation Big Data fait écho, hormis quelques exceptions. Il peut sembler curieux de les inclure dans cette revue sur le Big Data. Or la définition de Big Data, comme nous l'avons vu, ne se limite pas au volume de données, mais à leur nature (nominatives, multi-dimensionnelles, automatiquement collectées...) et l'utilisation qui en est faite (modélisation, prédiction). Ces études sont des exemples de l'incorporation du Big Data sous ces différentes facettes :

- par l'exploitation de nouvelles sources de données : smartphones, réseaux sociaux, questionnaires en ligne massivement diffusés...
- par l'intégration de données à haute-dimension et multi-modales de manière globale
- par la mise en œuvre de techniques d'apprentissage automatique, de NLP, de modélisation de réseaux, et de visualisation des données
- dans certains cas, par l'accès à des bases de données ou des échantillons élargis.

On a davantage de nouvelles données, mais on fait aussi davantage parler les données déjà existantes.

Néanmoins, ces études exposent aussi les faiblesses de cette approche, en général, et plus particulièrement en psychiatrie. Il y a de nombreux points à disséquer dans chaque étude individuelle, mais on se limitera aux plus saillants. La première critique est méthodologique. Nombre de ces études ont des échantillons de faible taille, parfois seulement quelques dizaines de sujets, tout en examinant un grand nombre de variables. Or comme nous l'avons vu dans la brève présentation de l'apprentissage statistique (I.3.b., p. 44), la haute-dimensionalité soulève de nombreux problèmes d'inférence et de fiabilité. Même avec des techniques de réduction de dimensions et de sélection/régularisation de variables, il y a un très grand risque de surapprentissage, et donc de trouver une fausse association, avec de faibles effectifs. Ce risque est exacerbé, et surtout invérifiable, en l'absence de validation sur un échantillon externe/indépendant. La plupart des études présentées ici, qui mettent en œuvre des techniques d'apprentissage statistique automatique, utilisent des techniques de validation croisée (*cross-validation*) pour construire, sélectionner et ajuster leur modèle statistique. Mais sans un test du modèle sur un échantillon complètement indépendant de l'échantillon de départ, il est impossible de déterminer si le modèle est fiable et généralisable. Or une telle validation externe rigoureuse n'est observée que dans quatre études ici (Dosenbach et al., 2010; Gillan et al., 2016; Xiong et al., 2015 ; et avec certaines réserves Just et al., 2017),

ce qui pose de sérieuses limites aux conclusions qui peuvent être tirées des autres études.

Le choix du type de sujets contrôles, les biais dans la constitution des échantillons et le contrôle inadéquat de facteurs confondants restreignent encore davantage l'intérêt clinique des résultats. Par exemple, dans l'étude de Just et al. sur le langage et le suicide, une comparaison avec des personnes dépressives non-suicidaires aurait pu être plus pertinente, compte tenu des perturbations thymiques chez leurs sujets suicidaires. Dans un autre exemple, l'étude de Gillan et al. (cognition et compulsion, Gillan et al., 2016), et celle de Wu et al. (cognition et phénotypes bipolaires, Wu et al., 2017) présentent des facteurs confondants significatifs pour des études sur la cognition, et des déséquilibres dans leurs échantillons : l'âge et le quotient intellectuel que Gillan et al. tentent de contrôler dans leur modèle, et le niveau d'éducation chez Wu et al. qui n'est pas contrôlé. On peut accorder que ce manque de rigueur fasse partie des difficultés rencontrées dans la réalisation d'études préliminaires, ou études « preuve de concept » (*proof-of-concept studies*), et elles sont souvent abordées ouvertement par les auteurs, mais pas toujours de manière suffisante ou satisfaisante. Sans une certaine confiance en la fiabilité des études, ou une discussion éclairée de leurs limites, on se demande que révèlent au juste ces données et ces résultats, quelle est leur contribution à la connaissance clinique et scientifique.

Cette interrogation correspond à la deuxième critique, plus épistémologique. Le halo d'objectivité qui entoure les données quantifiables occulte le nombre de choix subjectifs qui doivent être faits dans la collecte, l'analyse et l'interprétation de ces données. Ces choix, soutenus par une justification et une argumentation de qualité très variable selon les études, se voient dans des études aussi disparates que l'application CrossCheck (fréquence d'enregistrement, type de données choisies pour représenter le comportement des patients), ou l'outil SPANR (seuil de variation d'épissage, polymorphismes comparés, Ben-Zeev et al., 2017; Xiong et al., 2015). Cependant, le langage, parfois obscur et parfois excessivement confiant, des articles, le format rigide des publications qui s'adapte mal à la complexité des études, tendent à noyer ces subtilités. La lecture de ces papiers semble impliquer une acceptation implicite de la supériorité des données et des résultats présentés en vertu du simple fait qu'ils soient mesurables/quantifiables, comme s'il était présupposé que les données et résultats quantifiables sont toujours objectifs, et que les données objectives sont supérieures aux données subjectives, plus propices à mener à de nouvelles connaissances fiables, à des

certitudes. Or la supériorité des données utilisées ici (en qualité, en fiabilité) est loin d'être établie, comme le montre la critique précédente. En réalité, ces données et leur analyse contiennent un degré non négligeable d'incertitudes, d'approximations et de subjectivité. Et la surabondance embarrassante de données, comme celles provenant de l'application CrossCheck ou des réseaux sociaux, peut devenir difficile à analyser et interpréter. Le défi face à ce flot de données est d'y trouver du sens sans y imposer un sens.

Cela ne veut pas dire qu'une approche Big Data soit inutile, mais plutôt qu'il faille approcher ces études avec un œil critique quant à l'origine et les limites des données et des résultats, tâche difficile sans une compréhension adéquate des méthodes en jeu et une explication claire et adaptée de leur utilisation dans ces études. Cet esprit critique doit s'étendre à l'idée d'objectivité et au leurre du tout quantifiable, qui peuvent nous aveugler sur la complexité et les incertitudes sous-jacentes. Plus fondamentalement, les interrogations que soulève le Big Data semblent aller au cœur du savoir, et à la manière dont sa définition, sa création et sa valeur sont modifiées par le Big Data.

2. Classifier

Compte tenu des débats autour des catégories diagnostiques, la recherche de nouveaux outils pour caractériser et nommer avec certitude les troubles des patients suscite beaucoup d'intérêt. Comme nous l'avons vu, les tâches de classification à partir de données complexes sont une des applications phares des techniques d'apprentissage automatique (voir I.3.b. p. 46). Dans le domaine psychiatrique, ces techniques peuvent servir au diagnostic, ou à la « stratification » des patients, c'est-à-dire leur ségrégation en sous-groupes phénotypiquement homogènes. De multiples études les mettent en œuvre en exploitant différentes sources de données, allant des dossiers médicaux informatisés, à l'imagerie. Certaines études opèrent au sein de cadres nosologiques connus (DSM, CIM), recherchant un « biomarqueur » permettant d'assigner ou de réfuter un diagnostic, ou essayant de raffiner les catégories diagnostiques par l'identification et la classification de sous-groupes. D'autres études cherchent à réinventer un système de classification en abordant les symptômes et d'autres données de manière transversale. L'objectif final peut être clinique – dépister des troubles sous-diagnostiqués, mieux distinguer des tableaux cliniques ambigus ou incertains comme un premier épisode dépressif bipolaire d'un épisode unipolaire – mais il peut aussi être orienté vers la recherche – construire des

cohortes de patients plus homogènes pour augmenter la puissance d'études génétiques ou pharmacologiques, par exemple.

a. Dossiers médicaux électroniques

Comme avant, le Big Data se retrouve ici à la fois dans la complexité des données (imagerie, texte libre, multi-modalité etc) et dans la quantité des données, comme dans les études utilisant des dépôts de dossiers médicaux, ou de larges cohortes en génétique. Dans une étude utilisant environ 5000 dossiers de patients ayant fréquenté un hôpital psychiatrique danois de 1998 à 2008, Roque et al. appliquent des techniques de NLP pour extraire des informations cliniques à partir du texte libre et structuré des dossiers, permettant d'enrichir le simple codage administratif qui est souvent insuffisant, incomplet et imprécis pour résumer l'histoire clinique du patient (Roque et al., 2011). Les informations extraites servent à associer des codes/diagnostics CIM-10 aux patients, sur la base d'un dictionnaire de termes cliniques rapportés aux codes CIM. Les auteurs utilisent ensuite ces codes pour trouver des comorbidités d'intérêt en identifiant les paires de codes qui apparaissent ensemble fréquemment et en leur recherchant des bases génétiques communes dans un catalogue de gènes (OMIM) ; puis pour distinguer des sous-groupes de patients par la technique du regroupement hiérarchique. Cette extraction et extrapolation automatique d'informations retrouve en moyenne 12 codes par patient, dont une majorité correspond au chapitre V de la CIM (troubles mentaux), avec un tiers des codes provenant du chapitre XVIII (symptômes, signes et résultats anormaux d'examens cliniques et de laboratoire) reflétant la prise en charge de plaintes somatiques plus générales. La recherche de comorbidités retrouve 802 paires de codes corrélés. Mais la plupart de ces corrélations sont soit des comorbidités déjà bien connues, soit des artefacts liés à la non-indépendance statistique des codes CIM, et à l'organisation particulière du centre hospitalier en question et son offre de soins spécialisés. En parcourant les catalogues de gènes à la recherche de nouveaux réseaux géniques et fonctionnels communs entre les troubles corrélés, les auteurs découvrent un lien possible, et assez surprenant, entre l'alopécie et la migraine. Le regroupement hiérarchique, servant à identifier des sous-groupes ou clusters de patients ayant un profil similaire, retrouve une stratification, au sein du groupe de patients codés « schizophrénie », selon le type de substance psycho-active consommée (alcool, cocaïne, opiacés etc), et un autre

regroupement de patients présentant des troubles psychiques liés à une lésion cérébrale, ainsi que d'autres clusters associés à des troubles somatiques. Cette classification particulière semble refléter, encore une fois, le type de soins proposés au site de l'étude, et donc la sélection du type de patient qui y est traité. Hormis l'intérêt des outils d'extraction d'information à partir de dossiers médicaux pour automatiser l'analyse des dossiers et l'exploitation de cette mine de données, cette étude souligne surtout à quel point l'organisation, et le contexte de recueil des données détermine les schémas, les liens et les corrélations qui apparaîtront dans l'analyse, surtout non-supervisée.

Un autre exemple d'extraction d'information est présenté dans l'étude de Jackson et al., basée sur la banque CRIS, qui regroupe tous les dossiers médicaux anonymisés du South London and Maudsley NHS Foundation Trust, l'un des plus larges groupes de santé mentale en Europe (Jackson et al., 2017). Les auteurs y développent un modèle SVM de classification de phrases pour détecter les références à des symptômes psychiatriques notables, dans les dossiers médicaux. Ils utilisent un lexique sémiologique, compilé par des psychiatres associés à l'étude, centré sur 50 symptômes présents dans les troubles mentaux « sévères » (schizophrénie, trouble schizoaffectif, trouble bipolaire uniquement). Les symptômes appartiennent à cinq grandes catégories détaillées dans l'étude : symptômes positifs, négatifs, maniaques, syndrome dissociatif, catatonie. Pour chaque symptôme, un modèle de classification binaire SVM est construit et entraîné par validation croisée sur une série de dossiers de patients. La précision de chaque modèle est ensuite vérifiée sur une autre série de dossiers « gold-standard » pour chaque symptôme, annotés, indépendants des dossiers d'entraînement. Ces dossiers proviennent d'une cohorte de 18 761 patients avec « trouble sévère », et 57 999 patients avec un diagnostic autre (cela inclut les patients ayant un diagnostic de dépression unipolaire, trouble anxieux, trouble de la personnalité etc). Quatre modèles (perte de motivation, baisse de l'humeur, stéréotypies, diffidence) sont exclus pour mauvaise performance. Puis les modèles restants sont appliqués aux comptes-rendu d'hospitalisation de 15 537 patients (7962 sévères, 7575 autres). Ils sont capables d'extraire des symptômes chez 87 % des patients sévères, et 60 % des patients autres. Les auteurs observent que la catégorie de symptômes la plus communément retrouvée correspond aux symptômes positifs, mais notent que les modèles sous-estiment probablement la présence de symptômes. Surtout, ils remarquent la prévalence importante de symptômes chez les patients avec un diagnostic autre, y compris des symptômes comme les hallucinations ou la paranoïa,

indiquant leur non-spécificité diagnostique, et leur transversalité à travers les catégories diagnostiques.

Plutôt qu'un outil de classification des patients selon les nosographies existantes, cette étude décrit un outil d'extraction automatique d'informations sémiologiques pour enrichir le diagnostic catégoriel. Comme l'exemple précédent, son application est, à ce stade, plus pertinente à des fins de recherche (constitution et suivi de cohortes par exemple), comme possible substitut à des questionnaires ou entretiens structurés. Ce qui ressort de la description de cet outil est la quantité de travail et de ressources nécessaires en amont – nettoyage des données, constitution d'un corpus de textes annotés par des experts, etc – afin de pouvoir exploiter ces données massives pour une problématique donnée. La liste de symptômes modélisés, compilée manuellement sur la base de leur importance en recherche et de leur fiabilité inter-observateur, est également guidée par cette problématique de recherche, et ne se prête pas forcément aussi bien à des applications cliniques par exemple. La question posée reste, en effet, primordiale pour guider l'utilisation des données et le choix des outils d'analyse. En soi, l'accès à des données massives ne supprime pas la nécessité du travail sur la problématique et sur la qualité des données, et ne garantit pas un résultat (information, modèle, outil...) systématiquement généralisable.

L'étude de Castro et al. décrit une application concrète d'extraction de données à partir de dossiers médicaux, avec classification diagnostique, aboutissant à la constitution d'une cohorte de sujets pour une étude génétique (Castro et al., 2015). Pour le bénéfice d'un des sites (Massachusetts General Hospital) du consortium International Cohort Collection for Bipolar Disorder, les auteurs minent les dossiers d'un groupement hospitalier couvrant 4,2 millions de patients depuis plus de 20 ans, à la recherche de sujets bipolaires et de sujets sains à inclure. Après un tri initial de dossiers selon leur codage administratif (donnant 52 235 patients), 209 sont sélectionnés au hasard comme base d'entraînement pour un algorithme d'apprentissage automatique. Ces dossiers sont relus par trois psychiatres afin de confirmer le diagnostic et de compiler une liste de termes ou d'attributs (*features*) contribuant au diagnostic. Des techniques de NLP extraient ces termes/attributs du texte des dossiers, pour les rendre accessibles à l'algorithme. À partir d'une liste de 414 attributs, la technique de réduction de variables *lasso* sélectionne 13 attributs pertinents que le modèle, une régression logistique, utilise pour apprendre à classer les dossiers d'entraînement en bipolaire/non-bipolaire. Une fois l'algorithme développé, il est déployé sur le reste des dossiers triés, où il identifie 1776 cas de patients bipolaires. Des

modèles de régression supplémentaires sont développés pour distinguer des sous-phénotypes d'intérêt (age de début, sous-type, antécédent familial de bipolarité, psychose etc).

En parallèle, pour pouvoir prendre en compte les dossiers plus anciens, partiellement numérisés, les auteurs utilisent également plusieurs algorithmes logiques (*rule-based algorithm*) qui classifient les dossiers selon un certain nombre de règles fixes applicables sur les seules données administratives (fréquence d'apparition du diagnostic codé, présence/absence d'autres diagnostics, présence/absence de traitements spécifiques au trouble bipolaire etc). Les auteurs identifient ainsi plus de 19 000 cas supplémentaires, ainsi que des sujets témoins. En comparant, chez 190 des sujets, le diagnostic attribué par les algorithmes, au diagnostic attribué en aveugle par un clinicien après un entretien SCID (Structured Clinical Interview for DSM-IV), les auteurs estiment la valeur prédictive positive des deux types d'algorithmes, qui se situe à 0,86 pour l'apprentissage automatique, et à 0,84 pour l'algorithme logique utilisant les règles les plus strictes. L'exploitation de cette large base de dossiers médicaux, par une combinaison de NLP, d'apprentissage automatique, d'algorithmes logiques, permet ici une classification des patients suffisamment précise pour automatiser, en recherche génétique, la constitution d'une cohorte de plusieurs milliers de patients (4500) et de sujets témoins (5000). Il est important de noter que, malgré la distinction entre les données d'entraînement et les données de validation, les auteurs n'ont pas pratiqué une validation externe sur un échantillon provenant d'un autre groupement hospitalier ou d'un autre site, par exemple. On ne sait donc pas si ces outils sont généralisables, d'autant plus que l'échantillon de validation de 190 sujets est petit. On remarquera, de plus, que les algorithmes logiques, notamment ceux à critères stricts, ont une performance comparable au modèle par apprentissage automatique, alors qu'ils sont basés sur une quantité moindre de données. Ceci renvoie à la citation rencontrée au chapitre précédent, mettant en garde contre les désavantages des données massives d'un point de vue statistique⁹, idée qui se manifeste aussi dans la réduction du nombre d'attributs pertinents de 414 à 13, pour le modèle par apprentissage automatique.

9 « Thus, we see that new technologies that allow for the collection of measurements for thousands or millions of features are a double-edged sword: they can lead to improved predictive models if these features are in fact relevant to the problem at hand, but will lead to worse results if the features are not relevant. Even if they are relevant, the variance incurred in fitting their coefficients may outweigh the reduction in bias that they bring. » (James et al., 2013)

b. Imagerie

Ces travaux sur les dossiers médicaux informatisés cherchent à automatiser la classification de patients ayant déjà des éléments de diagnostic. D'autres études recherchent plutôt des biomarqueurs permettant d'établir un diagnostic. La majorité des études utilise l'imagerie cérébrale (morphologique, fonctionnelle, de diffusion), mais certaines se basent sur l'EEG (électroencéphalogramme), des échelles de symptômes et d'autres questionnaires, des explorations génétiques, des tests neuro-cognitifs, des marqueurs sériques, ou une combinaison de données multimodales. On peut d'emblée noter qu'il est difficile de tirer des conclusions générales à partir de ces études, précisément parce que la plupart d'entre elles ne sont pas généralisables. Elles reposent sur de faibles effectifs (souvent moins de 1000 sujets, parfois moins de 100), sans validation externe, ni homogénéité entre les variables utilisées. Dans une revue de la littérature sur l'imagerie et l'apprentissage automatique dans la dépression, Patel et al. identifient onze études sur le diagnostic (Patel et al., 2016), avec des effectifs allant de 19 à 40 sujets par groupe, utilisant des modalités d'imagerie différentes (IRM fonctionnelle avec signal BOLD cerveau entier, ou connectivité entre régions, IRM morphologique en séquence T1 avec analyse du cerveau entier ou substance grise seule). On observe une prédilection pour les classificateurs SVM, certaines études choisissant néanmoins de tester plusieurs types de modèles en même temps. La plupart utilisent des techniques de réduction/sélection de variables, mais diffèrent dans le choix du type de technique. Dans la majorité des études, la construction/validation interne du modèle est faite par validation croisée, et plus précisément par *leave-one-out cross-validation* (LOOCV), méthode la plus adaptée aux faibles effectifs, mais susceptible à une variance importante. La précision de la classification diagnostique, dans les modèles développés par ces études, varie de 72 % à 94 %, reflétant la variabilité méthodologique des études. Mais sans validation externe, les résultats élevés signent probablement un sur-apprentissage plutôt qu'un modèle fiable performant.

La revue de Librenza-Garcia et al. (Librenza-Garcia et al., 2017) sur l'apprentissage automatique dans la recherche sur le trouble bipolaire, retrouve les mêmes limites que la revue de Patel et al., à savoir le manque de validation externe et le faible nombre de sujets. Ils identifient 38 études sur la classification diagnostique, dont onze en imagerie

morphologique et treize en imagerie fonctionnelle, cinq études génétiques, quatre en EEG, trois utilisant des tests neuropsychologiques, et deux utilisant des marqueurs sériques. Ici encore, l'hétérogénéité des études empêche de formuler des conclusions générales, mais les auteurs font quelques observations importantes :

- les études en imagerie sont souvent réalisées sur des patients ayant déjà un diagnostic et plusieurs années de traitement médicamenteux, limitant probablement l'intérêt de ces études pour les phases initiales du diagnostic ou du processus pathologique
- en revanche, les travaux impliquant des patients aux stades précoces de la maladie semblent montrer qu'il est plus difficile de distinguer les patients des sujets contrôles en début de maladie (Mwangi et al., 2016), précisément au moment où un outil d'aide au diagnostic aurait plus d'intérêt
- les outils de réduction/sélection de variables amélioreraient la précision et la performance du modèle de classification, notamment en imagerie, soutenant l'importance d'un modèle parcimonieux
- les études comprenant une validation externe montrent une baisse non-négligeable de la performance du classificateur sur l'échantillon externe, soulignant la difficulté à construire un modèle généralisable, et l'importance de la validation externe
- enfin, les auteurs remarquent qu'aucune des études citées ne propose un modèle plus performant que le système diagnostique actuel.

En pratiquant une méta-analyse sur 12 études de cette revue, comparant des sujets sains aux patients bipolaires en imagerie, les auteurs retrouvent une dispersion importante des études autour de la courbe SROC (*summary ROC*), qui a une forme approchant la diagonale plutôt que les asymptotes (Figure 16), signalant l'hétérogénéité et une performance moyenne des modèles développés dans ces 12 études, malgré des résultats individuels qui peuvent paraître impressionnants.

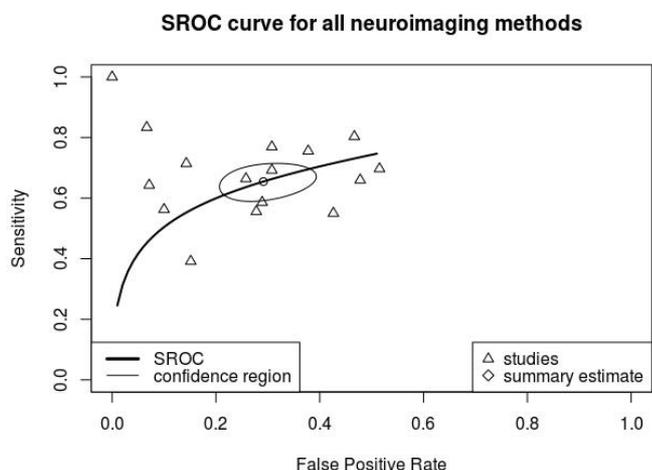


Figure 16 : Courbe sROC des résultats de méta-analyse des études de classification du trouble bipolaire en imagerie (Librenza-Garcia et al., 2017)

Zarogianni et al. tirent des conclusions semblables dans une revue synthétisant la recherche de biomarqueurs en imagerie dans la schizophrénie (Zarogianni, Moorhead, & Lawrie, 2013), notamment l'importance des techniques de sélection/réduction de variables et la moindre performance des modèles d'apprentissage automatique dans les stades précoces de la maladie (dans le cadre du premier épisode psychotique). Selon eux, les différences dans les protocoles de traitement et d'analyse d'images, en amont même du travail de classification, ainsi que les critères, plus ou moins stricts, de recrutement des sujets, seraient une source importante d'hétérogénéité entre les études, et des obstacles à la généralisation.

Deux études se distinguent dans cette revue par la présence d'une étape de validation sur un échantillon séparé, dont une avec une réelle validation externe, sur un groupe de sujets indépendant de la cohorte ayant servi au développement du modèle de classification. L'étude de Kawasaki et al. tente d'identifier la distribution de substance grise en IRM morphologique, en morphométrie basée sur les voxels (VBM), permettant de distinguer les patients schizophrènes des sujets non-malades (Kawasaki et al., 2007). Le modèle statistique (modèle linéaire multivarié) est entraîné sur un échantillon de 30 patients et 30 sujets sains, atteignant une précision de classification de 90 % sur cette cohorte. Les auteurs identifient les régions suivantes comme pertinentes à la classification : cortex préfrontal médial et latéral, insula, régions temporales latérales et médiales, et thalamus (diminution de substance grise chez les patients), et putamen et cervelet (augmentation de substance grise). Dans l'étape de validation, réalisée sur un échantillon test de 16 patients et 16 sujets sains, différent du premier groupe mais pas totalement indépendant (sujets recrutés dans le même centre notamment), le modèle atteint une précision de classification de 80 %.

Avec une cohorte d'entraînement beaucoup plus importante comprenant 239 sujets (128 patients schizophrènes, 111 sujets sains), l'étude de Nieuwenhuis et al. décrit le développement d'un classificateur SVM à partir des densités de substance grise (en VBM) sur les IRM morphologiques (Nieuwenhuis et al., 2012). Sur cette cohorte, le modèle départage les patients et les sujets sains avec une précision de 71,4 %. Testé sur un échantillon indépendant de 277 sujets (155 patients schizophrènes, 122 sujets sains) scannés sur une IRM différente du premier groupe, le même modèle atteint une précision de classification de 70,4 %, ce qui est en faveur d'un modèle stable et généralisable. Parmi les régions identifiées comme étant pertinentes à la classification, certaines, mais pas toutes, se recoupent avec l'étude précédente : les auteurs observent une diminution de

la densité de substance grise chez les patients dans les lobes frontaux, temporaux supérieurs et l'hippocampe, et une augmentation de la densité dans le lobe occipital gauche et les noyaux gris centraux.

Une étude plus récente, en imagerie de diffusion (plus précisément, diffusion spectrum imaging), utilise un protocole automatisé de standardisation et de traitement des images, développé par les auteurs, pour générer une matrice représentant l'intégralité et l'intégrité des faisceaux de substance blanche (Y. J. Chen et al., 2018). Pour identifier des schémas de substance blanche caractéristiques de schizophrénie, les auteurs comparent systématiquement, dans une cohorte de 108 patients schizophrènes et 144 sujets sains, les valeurs au sein de ces matrices générées pour chaque sujet, faisant le choix de cette technique plutôt que des méthodes d'apprentissage automatique, pour avoir un modèle plus explicable. Les règles de classification qu'ils dérivent de leurs comparaisons, permettent ensuite de classer une cohorte différente de 60 patients et 60 sujets sains, avec une précision de 76 % (AUC 0,79). Les auteurs notent une disparité entre les profils de substance blanche chez les hommes et les femmes, avec des modifications plus localisées chez les hommes et plus diffuses chez les femmes, se traduisant en une précision de classification de 85 % chez les hommes (AUC 0,87) et de 75 % chez les femmes (AUC 0,77). Par souci de lisibilité, les auteurs évitent l'apprentissage automatique, mais la prise en compte des données d'imagerie de diffusion dans leur ensemble, nécessite par conséquent le déploiement d'une multitude d'étapes complexes et parfois peu accessibles à la plupart des lecteurs.

En pratique clinique, face à une personne présentant déjà des symptômes, la difficulté résiderait davantage dans le choix d'un diagnostic psychiatrique, que dans la distinction entre malade et non-malade. L'étude de Schnack et al. tente d'aborder cette question avec une cohorte de 198 sujets, comprenant 66 patients schizophrènes (SZ), 66 patients bipolaires (BP) et 66 personnes non-malades (Schnack et al., 2014). Les auteurs font le choix de la technique SVM, construisant trois classificateurs binaires (SZ/HC, SZ/BP, BP/HC), sur les données des densités de substance grise en IRM morphologique. Ils obtiennent des modèles performants pour la classification des patients schizophrènes dans la cohorte d'entraînement (précision de 90 % pour SZ/HC et 88 % pour SZ/BP), mais bien moins performants pour distinguer les patients bipolaires des sujets sains (53 % de sensibilité et 67 % de spécificité, dans le modèle BP/HC). La validation externe des modèles se fait sur un échantillon indépendant de 46 patients schizophrènes, 47 patients bipolaires et 43 sujets sains, dont les IRM ont été capturées sur un appareil différent. La

performance des modèles diminue dans ce groupe test, avec des précisions de classification à 76 % (SZ/HC), 66 % (SZ/BP) et 61 % (BP/HC), et surtout des valeurs de sensibilité et spécificité déséquilibrées, avec par exemple, dans le modèle SZ/HC, une spécificité de 91 % et une sensibilité de 54 %. Les traits contribuant le plus, dans les modèles, à la distinction des patients schizophrènes sont la diminution de la densité de substance grise dans le cortex pré- et orbito-frontal, ainsi que temporal supérieur. Les différences chez les patients bipolaires apparaissent plus diffuses, indicatives d'une plus grande hétérogénéité clinique chez ces patients, selon les auteurs. Au delà des limites à la généralisation à d'autres populations, cette étude laisse apercevoir une autre limite de l'apprentissage automatique, à savoir la difficulté de transfert de l'apprentissage à différentes questions ou tâches. Un cadre d'apprentissage et une technique spécifiques peuvent générer un modèle très performant pour une question (distinguer un patient schizophrène d'un patient bipolaire ou d'un non-malade), mais échouer pour une problématique apparemment équivalente ou complémentaire (distinguer les patients bipolaires). Ceci met en relief la complexité du déploiement de ces outils.

Cette complexité est à l'œuvre dans l'étude de Drysdale et al., dont la rigueur méthodologique est à la mesure de la précision des résultats obtenus (Drysdale et al., 2017). À la recherche de biomarqueurs permettant d'élucider l'hétérogénéité des patients dépressifs et de prédire la réponse à la stimulation magnétique trans-crânienne (rTMS), les auteurs analysent les IRM fonctionnelles de repos réalisées chez une large cohorte multi-site de patients dépressifs résistants, pour identifier les schémas de connectivité cérébrale fonctionnelle distinguant des sous-groupes de patients, identifiant des réseaux avec 258 nœuds d'intérêt. Pour obtenir une représentation à dimensions réduites, ils recherchent au sein des réseaux de connectivité extraits pour chaque sujet, des ensembles de variables corrélées à la symptomatologie dépressive, obtenant ainsi deux représentations de connectivité associées à des combinaisons de symptômes, dénommées « anhedonia-related connectivity features » et « anxiety-related connectivity features ». L'étape suivante consiste à observer comment sont distribués les patients selon ces deux représentations (contenant une information clinique et fonctionnelle), s'il existe des clusters, ou sous-groupes. Les auteurs distinguent ainsi en regroupement hiérarchique, quatre sous-groupes ou « biotypes », dont ils explorent les caractéristiques. Puis ils développent des modèles SVM capables de classer les sujets selon leur diagnostic et leur biotype, à partir des données de connectivité les plus anormales, pour ensuite chercher si l'appartenance à un biotype prédit une meilleure réponse à la rTMS.

La recherche des schémas de connectivité, leur corrélation à une symptomatologie clinique, l'identification de clusters, et le développement des modèles de classification, sont entièrement réalisés sur une première cohorte de 711 sujets (378 contrôles de 12 sites, 333 patients de cinq sites, avec l'identification de clusters restreinte à 220 patients). Une validation externe des classificateurs est réalisée sur une cohorte différente de 477 sujets (125 patients de sept sites, 352 sujets sains de 13 sites), avec cinq sites uniques à cette cohorte. Les auteurs proposent aussi une validation longitudinale à quatre à six semaines, sur 48 à 50 sujets, et une validation de la spécificité sur 41 patients schizophrènes et 39 patients avec un trouble anxieux généralisé. Outre ces différents tests de généralisation, les auteurs mettent en œuvre plusieurs techniques et étapes d'optimisation : traitement des images et correction des artefacts ; estimation des schémas de connectivité par différentes approches ; réduction des dimensions ; test de différentes techniques d'apprentissage non-supervisé pour la recherche de clusters (regroupement hiérarchique et *k*-means clustering) ; sélection de variables pour la classification ; test de différents algorithmes de classification (régression logistique, SVM, LDA).

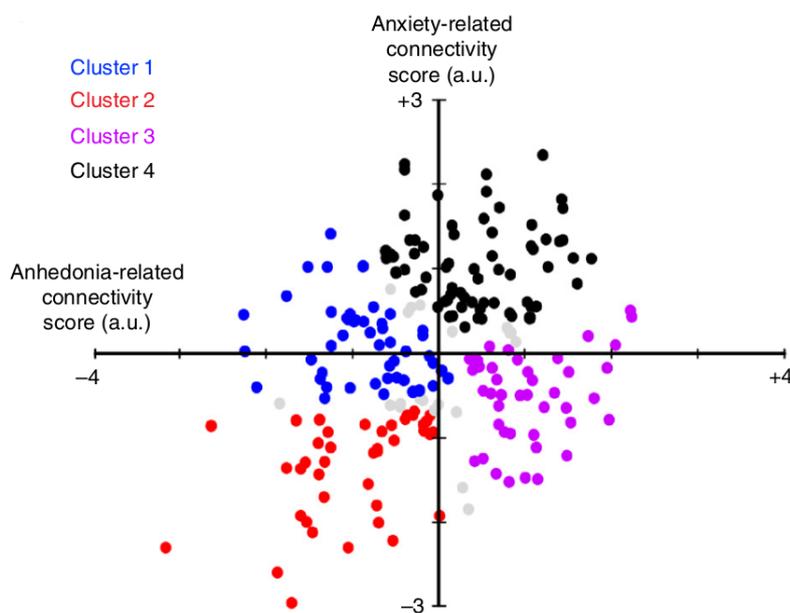


Figure 17 : Distribution de patients selon les dimensions *anxiety-related* et *anhedonia-related* des caractéristiques de connectivité fonctionnelle (Drysdale et al., 2017)

Les deux représentations de connectivité citées plus haut se distinguent par des régions cérébrales et une symptomatologie différentes. Le profil « anhedonia-related connectivity features », intégrant une représentation de la connectivité fronto-striatale et orbito-frontale, est corrélé aux symptômes d'anhédonie et de ralentissement psychomoteur, tandis que le profil « anxiety-related connectivity features », avec une représentation de la connectivité du système limbique (amygdale, hippocampe ventral,

striatum ventral, région pré-frontale latérale, gyrus cingulaire), est lié aux symptômes d'anxiété et d'insomnie. Les patients se distribuent en quatre sous-groupes selon les deux dimensions de ces représentations, comme les auteurs le visualisent sur la figure 17. La caractérisation plus précise des sous-groupes révèle un socle commun de connectivité anormale présent dans les quatre clusters, localisée dans l'insula, les régions orbito-frontale, pré-frontale ventro-médiale et sous-corticale. De même, trois symptômes sont présents chez la quasi-totalité des patients, indépendamment de leur sous-groupe, et ont une sévérité corrélée au degré de connectivité anormale : baisse de l'humeur, anhédonie et asthénie. Au delà de ces similitudes, des schémas distincts de connectivité anormale caractérisent les différents biotypes, qui semblent ici stratifier les patients de manière plus stable qu'une stratification basée sur leur symptomatologie. Les modèles classifiant les patients selon ces biotypes atteignent globalement une précision de 89 % sur la cohorte d'entraînement (sensibilité allant de 84-90 %, spécificité de 84 à 92%). Parmi 50 patients encore dépressifs, scannés une deuxième fois après quatre à six semaines, 90 % sont assignés à leur biotype de départ, suggérant une stabilité longitudinale. La validation externe de ces modèles sur une deuxième cohorte atteint une précision moyenne de 86 %, avec les meilleures performances (précision > 90 %) montrées par les classificateurs des biotypes 3 et 4, qui sont marqués par des niveaux plus sévères d'anhédonie. Testant la spécificité de ces biotypes à la dépression, les auteurs trouvent que 69 % des patients d'une cohorte ayant un diagnostic de trouble anxieux généralisé, sont assignés à un biotype dépressif par les classificateurs, 59 % parmi eux au biotype 4 qui est davantage associé à la dimension anxieuse (insomnie et anxiété plus sévères, anomalies de connectivité du profil anxiété). Ces patients anxieux, classifiés dans un biotype selon leur données de connectivité cérébrale, mais sans diagnostic clinique de dépression, semblent avoir des symptômes dépressifs plus sévères que les patients anxieux non-assignés. À l'inverse, seuls 9.8 % d'une cohorte de patients schizophrènes se voient assignés un des biotypes de dépression, suggérant peut-être des mécanismes fonctionnels et des anomalies de connectivité très différents entre dépression et schizophrénie, mais plus semblables entre anxiété et dépression.

L'attention portée à une méthodologie rigoureuse dans cette étude, donne des résultats qui montrent une certaine robustesse et une fiabilité, ayant à la fois un intérêt physiopathologique – dans l'identification sur le cerveau entier des anomalies de connectivité – et clinique, dans l'utilisation qu'en font les auteurs pour étudier une nouvelle cible de la rTMS et prédire la réponse à la rTMS (voir sous-chapitre suivant II.3). Certains

résultats, qui ne surprendront aucun clinicien au contact quotidien de patients, tels que la triade symptomatologique d'asthénie, anhédonie et baisse de l'humeur, ainsi que le recoupement avec l'anxiété généralisée, sont rassurants en quelque sorte : les résultats de cette étude ne sont pas surprenants et novateurs au point de devenir hors-sujet ou sans relation à la clinique ; et de la même manière, on voit que malgré la remise en question de la validité des cadres nosologiques formels, certains schémas pragmatiques et synthétiques que l'on adopte avec l'expérience clinique, restent tout à fait pertinents. Mais il y a aussi une certaine déception face à cette étude. Son apport à une compréhension des mécanismes psycho- et physiopathologiques semble minime, malgré l'approche en connectivité fonctionnelle. La stratification proposée, telle qu'elle est donnée, est difficile à se représenter de manière vivante et concrète. Ce problème n'est pas résolu par la réduction des biotypes aux deux dimensions « anxiety-related » et « anhedonia-related », dont le libellé paraît excessivement simplifié et au final, peu informatif. Enfin, l'impact clinique reste restreint, les modèles de classification étant difficilement intégrables dans une approche diagnostique de première ligne, l'étude ayant été conçue autour de la dépression résistante et des candidats à la rTMS.

Ceci évoque la critique émise par Woo et al. dans leur revue sur les biomarqueurs neuro-psychiatriques et la recherche translationnelle en imagerie (Woo et al., 2017) :

« Despite its great promise, neuroimaging has yet to substantially impact clinical practice and public health » (p. 365)

et plus loin : « major advances will require specific ways of combining pattern recognition and aggregated neuroimaging data that are not yet the norm » (p. 375).

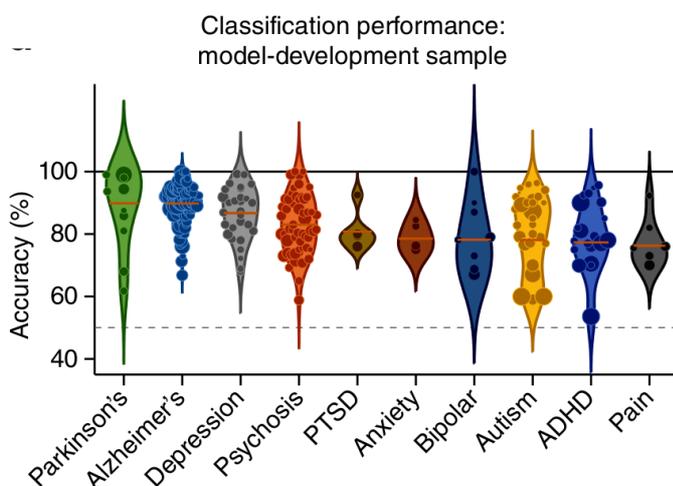


Figure 18 : Performance de différents modèles de classification de pathologies neurologiques et psychiatriques sur les données d'entraînement (validation interne) (Woo, Chang, Lindquist, & Wager, 2017)

Notant l'augmentation des études employant l'apprentissage automatique en neuro-imagerie, et des effectifs de leurs cohortes, surtout depuis 2013, les auteurs soulignent que la performance des modèles varie beaucoup entre les études, comme on peut le voir sur un résumé de leurs niveaux de précision sur la figure 18, et que seuls 9 % des modèles développés sont validés de manière prospective sur une cohorte indépendante de la cohorte d'entraînement. Ils mettent en garde contre le manque de lisibilité des modèles, rappelant que si l'on ne comprend pas comment un modèle a réussi, on ne saura pas non plus quand il sera susceptible d'échouer. Dans cette critique, ils insistent également sur l'importance d'un modèle plausible, c'est-à-dire ayant une certaine cohérence physiologique, et apportant une contribution scientifique à notre connaissance. En effet, plusieurs exemples de modèles performants mais spécieux donnent tout son sens à cette mise en garde, comme lors d'une compétition sur le diagnostic en imagerie de TDAH (trouble déficitaire de l'attention avec hyperactivité), où le modèle le plus précis semblait être principalement basé sur des signes de mouvement dans l'IRM, captant probablement un aspect comportemental bien connu du TDAH, sans apporter de nouvelles connaissances au niveau neurobiologique.

Les propositions de ces auteurs pour améliorer la recherche de biomarqueurs en imagerie, sont une réponse à ce paysage scientifique parsemé de publications *proof-of-concept*, mais avec une pénurie d'études prospectives, généralisables, et un manque d'intégration cohérente au sein du corpus de connaissances existant. Ils suggèrent, entre autres, de :

- focaliser les efforts de recherche sur des questions qui posent un défi ou qui échappent aux outils cliniques déjà existants, comme certains diagnostics différentiels, l'évaluation du pronostic et de la réponse aux traitements
- résumer et visualiser les modèles d'une manière accessible et facilement compréhensible aux lecteurs humains
- faire des liens entre les modèles et des résultats déjà connus dans la littérature
- accorder plus d'attention aux facteurs confondants
- utiliser des cohortes et séries de données plus « écologiquement » représentatives, à la fois par leurs caractéristiques démographiques et par le contexte de recueil de données
- développer des *process-based models* : plutôt que de chercher des biomarqueurs pour un symptôme spécifique ou un diagnostic particulier, ils proposent de chercher des signatures cérébrales associées à des fonctions intermédiaires (cognitives, sensori-motrices, émotionnelles...), dont la combinaison peut correspondre à un symptôme, et

qui peuvent être présentes dans une variété de symptômes et de syndromes. Cette approche permettrait de dépasser le biais des catégories diagnostiques actuelles, tout en offrant un cadre plus propice à la correspondance et à la traduction des connaissances anatomiques et physiologiques en applications cliniques.

Peut-être est-ce justement une attente irréaliste que d'exiger de ces études en imagerie, à la fois des outils de classification diagnostique performants et un apport significatif à la compréhension de la maladie, quand on sait que le compromis entre complexité et explicabilité, entre personnalisation et généralisation, et entre inférence et prédiction, est une tension centrale du *Big Data* et des outils utilisés pour l'appréhender. De plus, la part des anomalies cérébrales morphologiques et fonctionnelles dans la pathogenèse des troubles psychiatriques est encore mal élucidée, à la différence des pathologies neurologiques citées dans la revue de Woo et al. (maladie d'Alzheimer et maladie de Parkinson), qui sont au moins caractérisées par des lésions histologiquement et cytologiquement observables, servant de point d'ancrage diagnostique et théorique.

c. Génétique

L'héritabilité des maladies psychiatriques invite naturellement à l'exploration de la piste génétique, ce qui est l'objet des consortiums de GWAS (*genome-wide association study*), notamment le *Psychiatric Genome-Wide Association Consortium* (PGC), comparant les fréquences de SNPs chez des milliers d'individus pour trouver des facteurs de prédisposition génétique à l'échelle de populations (Sullivan, 2010). L'utilisation de ces données et de ces résultats pour un diagnostic ou un pronostic individuel passe par des scores de risque génomique individuels élaborés par différentes techniques (probabilités, apprentissage automatique/statistique, règles logiques). Plutôt que d'examiner la contribution, souvent faible, des polymorphismes à une maladie un à un et de manière indépendante, l'apprentissage automatique permet d'appréhender le génome de manière globale, prenant en compte l'ensemble des variations, leurs interactions et leurs effets non-linéaires, tout comme en imagerie où ces techniques permettent d'aborder les voxels de manière globale et interdépendante. Plusieurs études tentent cette approche en psychiatrie, pour la prédiction (abordée au prochain sous-chapitre) et la classification de patients.

Chuang et al. exploitent les données provenant de deux cohortes GWAS différentes, comprenant chacune plus de 2000 sujets (patients bipolaires et sujets sains), pour construire puis valider un modèle de risque génétique pour le trouble bipolaire (Chuang & Kuo, 2017). À partir de plus de 19 000 SNPs candidats identifiés dans la première cohorte comme étant liés au diagnostic de trouble bipolaire, les auteurs utilisent la méthode *random forest* (forêt d'arbres décisionnels, méthode d'ensemble non paramétrique) pour trouver l'ensemble de marqueurs le plus informatif ou significatif. Les marqueurs optimaux dans cet ensemble, raffinés par une méthode de sélection de variables (*stepwise selection*), servent à construire un classificateur par régression logistique multivariée. Dans la cohorte d'entraînement, l'ensemble des 289 SNPs identifiés par la *random forest*, distingue les patients des sujets sains avec une AUC de 0,944, tandis que le modèle de régression logistique, avec un ensemble de 121 SNPs, obtient une AUC de 0,924. Testée sur une cohorte de validation indépendante, l'efficacité de ces deux modèles baisse, atteignant respectivement une AUC de 0,702 (modèle à 289 SNPs) et 0,639 (modèle à 121 SNPs). Les auteurs construisent ensuite un score de risque génétique pour chaque sujet, en additionnant le nombre de marqueurs de risque, pondérés par leur coefficient de régression. Le score semble efficace pour distinguer les patients des sujets sains dans la cohorte d'entraînement (AUC de 0,9), mais n'en est pas capable dans la cohorte de validation (AUC de 0,5). Lorsque les auteurs inversent les cohortes dans les phases d'entraînement et de validation, la performance des modèles est comparable aux modèles précédents, indiquant une certaine stabilité de la méthode. Ils ne retrouvent, cependant, pas de SNPs en commun entre les modèles développés sur les deux cohortes, mais tout de même neuf gènes en commun (3,8 % des gènes identifiés) impliqués dans diverses fonctions cérébrales et dans la littérature sur le trouble bipolaire. En permettant la prise en compte d'une multitude d'effets génétiques mineurs et de leurs interactions, l'apprentissage automatique offre certes une avenue prometteuse pour la compréhension de la génétique des troubles psychiatriques, troubles hautement polygéniques avec des variations génétiques à effets faibles. Mais cela ne se traduira pas toujours aisément en un outil diagnostique simple, comme un score de risque, ou ne sera pas systématiquement supérieur à un score de risque polygénique déterminé de manière plus classique, par exemple, par l'identification des SNPs individuels les plus significatifs. L'hétérogénéité intra et inter-cohorte peut être un des freins majeurs, mais aussi le simple fait que les polymorphismes géniques sont seulement un facteur contributif parmi d'autres, et n'expliquent qu'une partie des variations pathologiques.

Pour cette raison, il peut être intéressant d'avoir une approche multi-modale qui combine les données génétiques avec des données d'une autre source. C'est ce que tentent Yang et al. dans leur méthode hybride fusionnant des données génétiques et d'IRMf pour détecter les patients schizophrènes (Yang, Liu, Sui, Pearlson, & Calhoun, 2010). Il s'agit d'une étude préliminaire sur seulement 40 sujets (20 patients, 20 sujets sains), sans groupe de validation externe, avec un array de SNPs limités à 384 SNPs provenant de 222 gènes, et une tâche d'IRMf auditive bien connue en recherche sur la schizophrénie (*auditory oddball task*). Après des étapes de sélection de variables, les auteurs construisent trois classificateurs SVM (basés sur les SNPs sélectionnés, les voxels sélectionnés, et une *independent component analysis* des voxels) dont la performance individuelle est légèrement inférieure à leur performance en tant qu'ensemble SVM (SNP : précision de 0,74 ; voxel : 0,82 ; *independent component* : 0,83 ; ensemble : 0,87).

À l'opposée de cette étude centrée très spécifiquement sur une pathologie, on trouve l'approche de Maier et al. qui tire profit des corrélations génétiques entre pathologies psychiatriques pour développer un modèle évaluant simultanément le risque de plusieurs troubles, dans l'espoir d'obtenir une évaluation du risque génétique plus performante (Maier et al., 2015). Partant de la cohorte du PGC, donnant accès à plusieurs milliers de sujets avec différents troubles psychiatriques (schizophrénie, trouble bipolaire, trouble dépressif majeur, trouble du spectre autistique et TDAH), les auteurs construisent un modèle linéaire mixte multivarié et appliquent la méthode de *multi-trait genomic best linear unbiased prediction*, leur permettant d'estimer un risque génétique global simultanément sur ces différents troubles. En comparaison à des modèles prenant en compte une seule pathologie à la fois (*single-trait genomic best linear unbiased prediction*), le modèle *multi-trait* est plus performant, dans la cohorte PGC d'entraînement, pour identifier les patients, surtout dans les cas de schizophrénie et de trouble bipolaire. Validé sur des cohortes indépendantes spécifiques pour la schizophrénie, le trouble bipolaire et la dépression, ce modèle *multi-trait* montre une performance significativement supérieure au modèle *single-trait* pour la classification des patients individuels, avec un gain supplémentaire en performance lorsque le modèle est limité aux SNPs associés à des gènes exprimés dans le système nerveux central. La supériorité de performance est d'autant plus marquée, dans les cas de schizophrénie et de trouble bipolaire, dans les déciles à plus haut risque génétique. Mais malgré ces gains en performance, la spécificité et la sensibilité du modèle restent trop faibles pour en faire un outil diagnostique individuel.

Le modèle montre un biais important (modèle trop éloigné des observations) dans les cas de dépression majeure, même s'il est statistiquement plus performant que le modèle *single-trait*. Et même pour la schizophrénie et le trouble bipolaire, la performance moyenne, correspondant au coefficient de corrélation entre le diagnostic clinique observé et le risque génomique individuel prédit, se situe entre 0,16 (trouble bipolaire) et 0,22 (schizophrénie). Le modèle est, comme le concèdent les auteurs, plus utile à la stratification de populations, qu'au diagnostic individuel.

Les difficultés inhérentes aux modèles basés sur des données génétiques, entre-aperçues dans les exemples précédents, sont clairement illustrées dans les études suivantes et les réponses qu'elles ont suscitées. À partir d'une cohorte néerlandaise pour l'étude de la dépression et de l'anxiété (plus de 3000 sujets), Milaneschi et al. proposent une stratification de la dépression selon trois sous-types obtenus par *latent class analysis*, technique de regroupement non-supervisée, sur les données sémiologiques (CHARGE inflammation working group et al., 2016). Ces sous-types sont nommés, assez maladroitement, « modéré », « sévère typique », caractérisé par une diminution du poids et de l'appétit, et « sévère atypique », caractérisé par une augmentation du poids et de l'appétit, les libellés ne correspondant pas tout à fait aux sous-types du même nom dans d'autres classifications. Après calcul de scores de risque génomique (*genomic profile risk score* ou GPRS) pour différentes pathologies à partir de données provenant de larges cohortes GWAS, les auteurs trouvent une association entre le score de risque psychiatrique et le diagnostic de dépression, mais pas d'association avec le score génomique métabolique (traits influençant l'indice de masse corporelle et la triglycéridémie). Plus spécifiquement, le GPRS schizophrénie explique une plus grande part du risque de dépression que le GPRS dépression, et le sous-type « typique » est fortement associé au GPRS schizophrénie, alors que le sous-type « atypique » est associé à la fois au score schizophrénie et au score métabolique, amenant les auteurs à conclure que leur stratification clinique aurait aussi un fondement génomique. La prédominance du score schizophrénie est probablement liée, comme l'indiquent les auteurs, à la plus grande stabilité de ce score due à la plus grande taille de la cohorte schizophrénie par rapport aux autres cohortes (et peut-être, pourrait-on proposer, une moindre hétérogénéité clinique et génétique). Ceci souligne à quel point l'interprétation et la comparaison de scores de risque génétique peut être hasardeuse. Mais plus fondamentalement, les sous-types de dépression retrouvés, et leur démarcation sur le critère de poids et d'appétit, peut poser question quant à leur validité clinique et

physiopathologique. Pour cause, deux réponses ont été adressées à cet article proposant que les catégories identifiées sont en fait des artefacts d'un modèle biaisé par des variables hautement corrélées, à savoir la baisse/augmentation du poids et de l'appétit (Mäkinen, 2018; van Loo, Wanders, Wardenaar, & Fried, 2018). Les fluctuations de variables plus précisément mesurables et hautement interdépendantes (plus il y a prise de poids, moins il y a perte de poids) sont plus « évidentes », et sont susceptibles de dominer un modèle issu d'une analyse non-supervisée, donnant des catégories artificielles sans lien étiologique à la dépression. Comme le suggèrent les réponses à l'article, un épisode dépressif pourrait tout simplement révéler, chez les sujets « atypiques », en les mettant en situation de stress, leur fragilité métabolique génétique se manifestant par une prise de poids. L'association des gènes du métabolisme au sous-type « atypique » n'est donc pas synonyme d'une contribution de ces gènes à l'étiologie de la dépression. Le choix des méthodes d'analyse, la bonne connaissance de leurs limites et des caractéristiques statistiques des données utilisées, avec les biais et les facteurs confondants possibles, sont donc absolument primordiaux pour une juste interprétation des résultats. Aussi doit-on insister sur la nécessité de bien maîtriser les données et leur niveau d'adaptation à la question posée, et de se prémunir contre une confiance excessive en la seule abondance de données, ou un modèle qui « marche ».

L'étude de Skafidas et al. propose un score de risque génétique pour les troubles du spectre autistique, basé sur 237 SNPs dont l'identification est guidée par l'analyse de voies géniques fonctionnelles (*gene pathway analysis*), le score étant développé par un modèle de régression sur une cohorte de plusieurs centaines de patients et sujets sains, puis validé sur une cohorte différente de patients combinée avec une troisième cohorte de sujets sains (Skafidas et al., 2014a). Les discussions et échanges de courriers au sujet de cette étude sont un exemple des difficultés de réplication et de généralisation, surtout en l'absence d'un modèle partagé de manière transparente et facilement accessible (Belgard, Jankovic, Lowe, & Geschwind, 2014; Robinson et al., 2014; Skafidas et al., 2014b, 2015). Elles mettent en évidence les disparités de résultats qui peuvent exister en fonction des méthodologies mises en œuvre, mais aussi les divergences, voire les erreurs, d'interprétation émanant de niveaux d'expertise/de maîtrise différents sur les données. En effet, à travers deux démonstrations séparées, Belgard et al. et Robinson et al. questionnent les résultats de Skafidas et al., suggérant que le pouvoir discriminant du score serait plutôt lié à l'hétérogénéité ancestrale entre les cohortes, et non à un lien étiologique à l'autisme. Les questions de biais, de facteurs confondants, d'hétérogénéité

entre les cohortes (ici la variabilité génétique entre groupes d'ascendance différente), sont très prégnantes en génétique, comme on le voit à travers ces exemples, mais concernent aussi bien tout autre domaine d'étude en bio-médecine, a fortiori dans l'utilisation de données massives et multidimensionnelles, de qualité variable. Ceci rappelle la réalité du travail d'analyste de données, avec sa part importante de nettoyage et de préparation des données, et en explique la nécessité. Il ne suffit pas d'avoir des données, que l'on n'a pas forcément collectées soi-même d'ailleurs, et des outils d'analyse, pour arriver à des conclusions intéressantes. Il faut aussi que ce jeu de données soit au minimum fiable (sans erreurs, à l'authenticité vérifiée) et utile à la question posée (de taille suffisante, comprenant des variables et des sujets pertinents, contrôlé pour les facteurs confondants etc). Pour s'assurer que ces critères soient remplis, une maîtrise et une connaissance intime et profonde des données est incontournable, et demande, en pratique, du temps et de l'expertise, donc des ressources importantes.

d. Biomarqueurs sériques

Les exemples présentés jusque-là sont peu applicables en clinique courante, comme le reconnaissent parfois les auteurs-même des études. En revanche, plusieurs études recherchant des profils sériques semblent plus proches de cet objectif. Deux d'entre elles (Chan et al., 2015; Haenisch et al., 2016) se distinguent par la taille des cohortes et les multiples étapes de validation. Chan et al. pratiquent un dosage immunologique multiplex (*multiplex immunoassay*) sur des échantillons de sérum recueillis auprès de cinq cohortes multi-centriques de patients schizophrènes et de sujets sains (127 patients, 204 contrôles au total), afin d'identifier, par méta-analyse, les molécules sériques qui varient entre patients et sujets sains. Ces marqueurs sont ensuite utilisés pour construire un modèle de régression logistique, avec sélection de variables par lasso, pour distinguer patients et sujets sains. Le panel d'une vingtaine de marqueurs sériques est validé sur deux cohortes (plus d'une centaine de sujets) indépendantes des cohortes de découverte, atteignant une AUC de 0,97 sur l'une des cohortes et une sensibilité de 89 % sur la deuxième. Les marqueurs identifiés sont impliqués dans l'inflammation, le système immunitaire, la signalisation hormonale et par facteurs de croissance, et l'hémostase. Prenant une approche rétrospective, les auteurs testent la capacité de ce panel à prédire les futurs diagnostics de schizophrénie, sur deux cohortes ayant eu des

échantillons de sérum prélevés avant la détermination d'un diagnostic (445 sujets au total) : une cohorte de l'armée américaine (*Department of Defense Serum Repository*, dépôt collectant les prélèvements obligatoires du personnel militaire), et une cohorte de personnes consultant pour symptômes psychiatriques sans pathologie psychiatrique établie, dont certaines étaient identifiées comme à très haut risque psychotique (*ultra-high risk*) par l'échelle CAARMS, *Comprehensive Assessment of At-Risk Mental State* (étude collaborative ICAAR). Le modèle, développé sur des patients déjà diagnostiqués, semble montrer une bonne efficacité dans la prédiction du diagnostic de schizophrénie dans ces cohortes (AUC de 0,9 dans la cohorte militaire et 0,82 dans la cohorte de demandeurs de soins). Dans la deuxième cohorte, l'inclusion du score CAARMS dans le modèle améliore sa performance par rapport à un modèle basé sur les biomarqueurs seuls ou le score CAARMS seul (AUC de 0,9 ; 0,82 ; 0,72 respectivement). Le modèle montre aussi une spécificité pour la schizophrénie, distinguant les sujets qui évolueront vers une schizophrénie de ceux qui évolueront vers un trouble bipolaire (AUC 0,91 dans la cohorte militaire), sans pouvoir identifier les sujets pré-bipolaires par rapport aux contrôles (AUC 0,53).

L'étude de Haenisch et al. a une approche très similaire pour la recherche de biomarqueurs sériques dans le trouble bipolaire (Haenisch et al., 2016), utilisant des outils statistiques similaires et le même schéma de méta-analyse des *multiplex immunoassays* (158 patients bipolaires et 143 sujets sains), suivie d'une validation externe (66 patients bipolaires, 44 sujets sains) et d'une recherche prospective sur des personnes non diagnostiquées au moment du prélèvement (même cohorte militaire que Chan et al. : 110 patients pré-bipolaires, 75 pré-schizophrènes, 184 sujets sains) ou présentant un premier épisode dépressif (cohorte néerlandaise NESDA pour la dépression : 90 patients dépressifs, 12 pré-bipolaires, 368 sujets sains). Les vingt protéines identifiées sont impliquées dans l'inflammation (sept marqueurs pro-inflammatoires, quatre marqueurs anti-inflammatoires), le transport lipidique et l'activité enzymatique des métallo-endopeptidases. En utilisant une sélection de ces marqueurs, mais avec des coefficients de régression recalculés sur la cohorte de validation (donc un nouveau modèle), le classificateur obtient une AUC de 0,92 sur la cohorte de validation. Dans l'étape prospective, le panel de marqueurs distingue douze patients secondairement diagnostiqués bipolaires, au sein d'une cohorte de 102 personnes ayant reçu un diagnostic de dépression après un premier épisode dépressif (AUC de 0,84), montrant un intérêt selon les auteurs, dans le cadre du diagnostic différentiel précoce de troubles

thymiques. Les marqueurs sont aussi capables de prédire un diagnostic de bipolarité chez des personnes pré-symptomatiques, par rapport à des sujets sains (AUC de 0,79), et des sujets qui seront ensuite diagnostiqués schizophrènes (AUC de 0,91). Ils semblent montrer une spécificité pour le trouble bipolaire, compte tenu de leur moindre performance ou leur inefficacité dans la classification et la prédiction de diagnostics de dépression (AUC de 0,64) et de schizophrénie (AUC de 0,52) par rapport aux sujets sains. Malgré des résultats intéressants, ces études ont des limites significatives. Mise à part l'hypothèse non-évidente qu'un prélèvement sanguin périphérique pourrait être un indicateur d'un processus pathologique possiblement localisé dans le système nerveux central, recalculer les coefficients de régression pour chaque cohorte (Haenisch et al., 2016) peut être une source majeure de surapprentissage et donc de surestimation de la performance, car il s'agit en pratique d'un nouveau modèle même si les marqueurs sont identiques. De plus, aucune des cohortes dans ces deux études n'est contrôlée pour le tabagisme, l'indice de masse corporelle ou l'utilisation d'un traitement hormonal ou contraceptif, comme le notent clairement les auteurs de Chan et al., citant l'absence d'information sur certaines cohortes. Au vu des marqueurs identifiés et des données disponibles sur le tabagisme pour certaines des cohortes, ces caractéristiques pourraient être des facteurs confondants significatifs, amenant à penser que les modèles seraient en train de détecter ces variations (fumeur/non-fumeur par exemple), plutôt qu'un facteur étiologique.

e. Approches multimodales

Entre le problème des résultats morcelés en imagerie et faiblement explicatifs en génétique, se trouve la limite d'un cadre exclusivement biologique (et individuel) pour comprendre un phénomène hautement multi-factoriel et lié à un contexte relationnel et environnemental. De plus, la recherche de biomarqueurs repose encore souvent sur ces mêmes catégories diagnostiques qui sont décriées pour leur manque de validité biologique, ce qui semblerait être à terme, un exercice futile. Les tentatives d'incorporer différentes modalités de données, que l'on a pu voir dans certaines de ces études, l'appel de Woo et al. pour une approche par *process-based models* (Woo et al., 2017), ainsi que toutes les études citées au sous-chapitre II.1., recherchant des manières plus originales de concevoir la pathologie psychiatrique ou de nouvelles sources de données, sont une reconnaissance de l'insuffisance d'un modèle purement biologique. Les études suivantes

étendent davantage l'approche multi-modale, en combinant une plus grande variété de données autres que biologiques (sémiologiques, cognitives, démographiques, autobiographiques etc), dans le but à la fois de détecter les patients, et de définir de nouvelles catégories.

Les études de Grisanzio et al. et Clementz et al. mettent cela en œuvre dans les domaines des troubles anxio-dépressifs (Grisanzio et al., 2018) et psychotiques (Clementz et al., 2015). À partir de données du consortium multicentrique B-SNIP (Bipolar-Schizophrenia Network on Intermediate Phenotypes), incluant 711 patients bipolaires, schizophrènes et schizo-affectifs, 883 parents du premier degré, et 278 sujets sains, Clementz et al. tentent d'identifier des biotypes non-liés aux diagnostics psychiatriques, mais uniquement déterminés par des mesures cognitives et électrophysiologiques (Clementz et al., 2015). Pour ce faire, ils se basent sur l'administration d'une batterie neuropsychologique (BACS : Brief Assessment of Cognition in Schizophrenia), de tests des fonctions exécutives (tâches saccade et anti-saccade, tâche signal stop) et la recherche des potentiels évoqués auditifs (*paired stimuli* et *oddball evoked*). Parmi tous les résultats des tests, ceux qui diffèrent significativement entre patients et sujets sains sont identifiés, puis subissent une analyse par composante principale (PCA) qui en donne une représentation par une combinaison linéaire des résultats, permettant de réduire le nombre de dimensions, de limiter les redondances entre les différentes mesures et d'augmenter le pouvoir discriminant de ces représentations. Ces combinaisons sont ensuite utilisées comme variables dans un algorithme de regroupement/partitionnement, sur les patients uniquement (*k-means clustering*, voir I.3.b. p. 48), pour obtenir des groupes de patients distribués selon ces combinaisons de résultats neuropsychologiques et neurophysiologiques. Puis en employant des analyses discriminantes multivariées, les auteurs estiment la capacité de ces biotypes à classifier les patients selon d'autres données (imagerie, héritabilité, échelle de fonctionnement global et scores de symptômes).

À partir de huit combinaisons de mesures qu'ils regroupent en deux domaines, contrôle cognitif et réactivité sensori-motrice, les auteurs obtiennent trois biotypes qui ne se superposent pas aux diagnostics cliniques DSM de départ et comprennent un mélange de diagnostics DSM, même si le biotype 1 contient une majorité de patients schizophrènes (59 %) et le biotype 3 une majorité de patients bipolaires (44 %). Selon les auteurs, les trois biotypes semblent avoir une meilleure homogénéité intra-groupe sur les mesures cognitives, neurophysiologiques, d'imagerie et de fonctionnement global, en comparaison

aux groupes classés par diagnostic DSM, et permettent de mieux rendre compte des variations dans ces mesures et dans l'état clinique des parents du premier degré. Le biotype 1 (n=198) présente un déficit cognitif important, une hyporéactivité électrophysiologique, une diminution diffuse du volume de substance grise, une proportion élevée de parents symptomatiques, et une atteinte importante du fonctionnement psychosocial. Le biotype 2 (n=235) montre aussi un déficit cognitif et une diminution de substance grise, mais à un degré moindre. En revanche, ce biotype montre une hyperréactivité électrophysiologique. Le biotype 3 (n=278) est celui qui varie le moins des sujets sains, sans anomalies significatives du fonctionnement cognitif, une hyporéactivité très modeste, des modifications cérébrales plus modestes et localisées, une proportion nettement plus faible de parents symptomatiques et un meilleur fonctionnement psychosocial que les deux autres biotypes. Ces schémas sont retrouvés chez les parents du premier degré, à un degré moindre.

La validité des catégories, ou biotypes, proposés ne peut être établie sans reproduction dans une cohorte indépendante, et confirmation de leur stabilité dans le temps. Les auteurs évoquent des études de validation en cours, mais à ce jour, il ne semble pas y avoir de travaux publiés dans ce sens. De plus, ces biotypes sont basés sur une dimension pathologique restreinte, représentée par des tâches cognitives et des réponses physiologiques très ciblées, qui n'expliquent pas toute la pathologie. On peut regretter qu'il n'y ait pas eu une meilleure mise en correspondance de ces données avec des données plus précises sur la symptomatologie des sujets, afin de voir de possibles représentations cliniques, au-delà de l'étiquette diagnostique. Néanmoins, cette étude montre que les catégories de psychose telles qu'elles sont définies dans le DSM ont peu de fondements neurobiologiques, et par conséquent, ne constituent pas un cadre pertinent pour une recherche étiologique biologique, voire même psychopathologique.

Ceci n'est pas très surprenant quand on considère que les catégories du DSM-3 et 4 (utilisé dans cette étude) ont été construites avec une volonté explicite d'éviter toute hypothèse étiologique ou psychopathologique (Kupfer & Regier, 2011), et que la manière dont ces diagnostics sont posés dans le contexte de la recherche, avec un entretien dirigé ou semi-dirigé, aplatit en quelque sorte la caractérisation des patients. On peut donc comprendre cette étude comme une remise en question du cadre nosographique dominant, plutôt que comme la fondation d'une nouvelle nosographie, d'autant plus qu'une conception catégorielle de la pathologie psychiatrique, implicite dans cette étude, n'est pas un a priori irréfutable. D'ailleurs, chacun des choix méthodologiques dans cette étude est

issu d'une orientation théorique : conception catégorielle de la maladie allant avec le choix d'un algorithme de partitionnement en groupes, littérature sur les anomalies de potentiels évoqués auditifs dans la schizophrénie, focalisation sur le fonctionnement cognitif et notamment les fonctions exécutives. Bien qu'elle soit souvent bien argumentée et discutée par les auteurs, cette orientation n'en reste pas moins un choix subjectif et réfutable, parmi d'autres possibles.

L'approche de Grisanzio et al. est centrée, en revanche, sur la sémiologie, partant d'une cohorte de 420 personnes (220 sujets sains, 100 patients avec trouble dépressif majeur, 53 avec trouble panique, 47 avec syndrome de stress post-traumatique), sans exclusion des patients avec des comorbidités psychiatriques accompagnant leur diagnostic principal (Grisanzio et al., 2018). Un questionnaire d'auto-évaluation, le *Depression, Anxiety and Stress Scale* (version 21 ou DASS-21), est administré à tous les sujets, malades et non-malades. Une analyse par composante principale permet de simplifier les réponses individuelles en composantes de symptômes, dont les scores sont ensuite utilisés pour effectuer un regroupement hiérarchique (*hierarchical clustering*) des sujets. Les auteurs identifient trois composantes qui rendent compte de plus de 70 % de la variance de symptômes, et les nomment « anhedonia », « anxious arousal » et « tension ». Le regroupement selon ces trois composantes révèle six groupes ou clusters d'individus (Figure 19), libellés de manière suivante par les auteurs, en moyennant les composantes de symptômes qui les constituent : « normative mood » (n=180), « tension » (n=81), « anxious arousal » (n=55), « general anxiety » (n=38), « anhedonia » (n=29), « melancholia » (n=37). Les clusters comprennent un mélange de patients avec des diagnostics différents. On trouve par exemple des patients avec dépression ou stress post-traumatique à travers les six groupes. Le même protocole appliqué à une cohorte indépendante (n=381) retrouve globalement les mêmes composantes de symptômes (avec quelques différences dans les items précis contribuant aux composantes) et un regroupement similaire en six clusters.

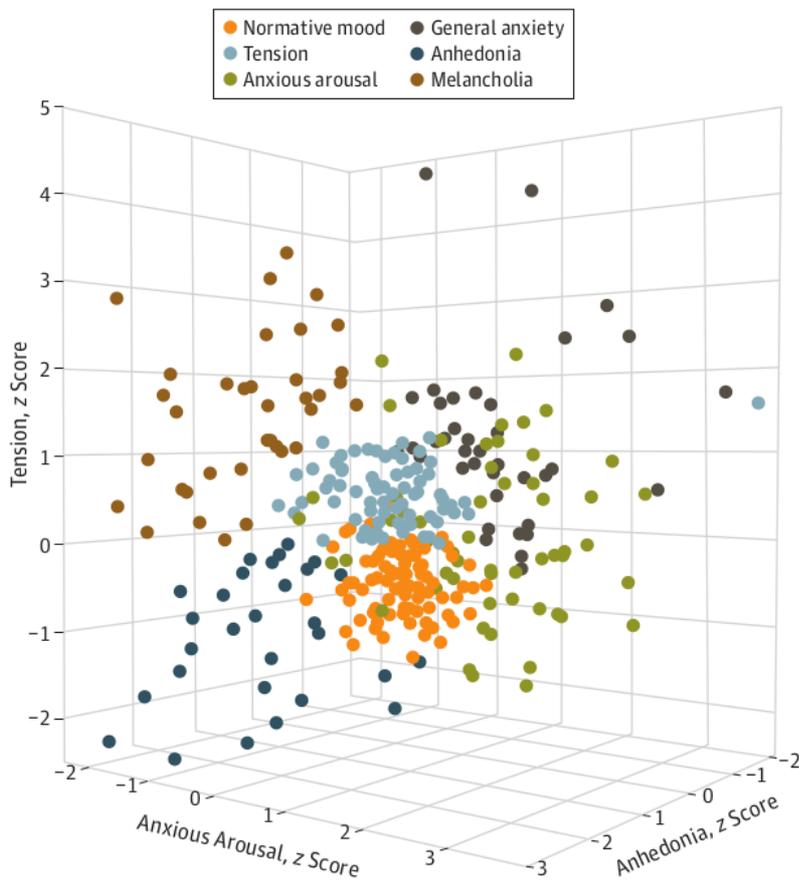


Figure 19 : Distribution de patients en 6 clusters selon les composantes *anhedonia*, *anxious arousal* et *tension* (Grisanzio et al., 2018)

Pour une meilleure caractérisation des six groupes, les fonctionnements cognitif (test IntegNeuro), électrophysiologique (EEG de repos et lors d'un stimulus émotionnel) et quotidien (*Brief Risk-Resilience Index for Screening*) sont évalués puis comparés entre les six clusters par ANOVA (*analysis of variance*). Les auteurs en déduisent que le cluster *tension* montre un fonctionnement préservé dans tous les domaines mesurés, alors que le cluster *anxious arousal* est caractérisé par une perturbation cognitive importante et un fonctionnement quotidien très atteint. Le cluster *general anxiety* a, en revanche, un fonctionnement quotidien préservé, mais une augmentation d'activité cérébrale pariéto-occipitale à l'EEG, face à un stimulus émotionnel. Une augmentation d'activité électrique cérébrale frontale au repos peut être observée dans le cluster *anhedonia*, tandis que le cluster *melancholia* a l'atteinte la plus sévère du fonctionnement quotidien, notamment dans le domaine social. Malgré la taille moyenne voire modeste des cohortes, l'un des intérêts de cette étude réside dans le choix de ne pas exclure des patients avec des comorbidités, se rapprochant davantage des situations rencontrées dans le monde « réel ». L'inclusion des sujets sains dans l'analyse de regroupement permet une ouverture sur une interprétation dimensionnelle plutôt que catégorielle des groupes. La

réplication, quoique imparfaite, dans une cohorte indépendante est encourageante, ainsi que l'utilisation d'un questionnaire d'auto-évaluation de symptômes comme mesure principale, ce qui offre au moins une vision du vécu et de l'expérience des patients, aussi incomplète et encadrée qu'elle puisse être. La question de la stabilité longitudinale de ce modèle demeure, tout comme celle de l'apport à la pratique clinique ou aux poursuites scientifiques.

Il est bon d'introduire ici l'étude de Whelan et al. qui s'inscrit au sein du large consortium européen IMAGEN, dont le but est d'étudier le développement et le comportement adolescents par une approche longitudinale neurobiologique, psychologique et sociale (Whelan, Watts, Orr, Althoff, Artiges, Banaschewski, Barker, Bokde, et al., 2014). Cette étude tente de distinguer des profils permettant d'identifier les adolescents en mésusage d'alcool, ou de prédire un risque de mésusage futur. Incluant au total 692 adolescents chez qui ils relèvent un grand nombre de données à l'âge de 14 ans, les auteurs explorent leur morphologie et leur fonctionnement cérébral (IRM et IRMf), leurs traits de personnalités (différentes échelles, dont certaines plus spécifiques aux problématiques d'addiction), leur fonctionnement cognitif (tests psychométriques, comme Wechsler Intelligence Scale for Children ou WISC, et fonctions exécutives), certains facteurs environnementaux (tabagisme et consommation d'alcool lors de la grossesse, statut socio-économique), une dimension génétique (gènes candidats) et une dimension autobiographique (*life experience*, évaluée par plusieurs questionnaires sur les expériences de vie potentiellement sources de stress ou de trauma). Ces traits étant potentiellement hautement corrélés, les auteurs utilisent une technique de sélection/réduction de variables (*Elastic Net regularization*), afin de minimiser les redondances et la variance dans les modèles de régression logistique qu'ils développent.

Le premier modèle distingue les adolescents en mésusage à l'âge de 14 ans (ayant eu au moins trois épisodes d'intoxication alcoolique aigüe, ou *binge drinking*), de leurs pairs sains (au maximum deux instances de consommation d'alcool jusqu'à l'âge de 16 ans au moins). Ce modèle, très performant dans la cohorte de découverte (AUC = 0,96), est cependant dominé par le tabagisme, qui coïncide très fréquemment avec la consommation d'alcool à cet âge, étant de ce fait, peu informatif sur la nature des autres facteurs de risque d'alcoolisation. Un second modèle, sans le facteur tabagisme, retrouve une bonne performance dans la cohorte de découverte (AUC = 0,90), qui se maintient dans une cohorte externe de validation avec une population phénotypiquement différente (AUC = 0,80). Cherchant les facteurs explicatifs les plus importants, les auteurs retirent un

à un les différents domaines de variables du modèle et évaluent la baisse de performance. Ils concluent ainsi que le facteur *History* (incluant les antécédents familiaux, le tabagisme ou l'alcoolisme maternel, les évènements de vie stressants) contribue le plus au modèle permettant d'identifier les adolescents en mésusage à 14 ans. Afin de trouver des facteurs de risque qui prédiraient un mésusage futur, les auteurs comparent les adolescents sains à des adolescents qui ne sont pas en mésusage à 14 ans mais le deviennent avant 16 ans. Le modèle, obtenu à partir des données relevées à 14 ans, montre de nouveau une contribution importante du domaine *History*, mais cette fois, les données sur la personnalité et les données d'imagerie cérébrale sont aussi très discriminants. L'application du modèle à une cohorte de validation externe indépendante soutient sa fiabilité (AUC = 0,75). La taille de la cohorte, le suivi longitudinal, avec une validation externe sur une cohorte indépendante et ayant des caractéristiques différentes, la variété des données incorporées, ainsi qu'une certaine attention portée à la méthodologie, la rédaction et la présentation détaillée, mesurée et intelligible, la mise en lien de multiples domaines de compréhension, ces éléments réunis dans cette étude en font un exemple de ce qu'une utilisation consciencieuse et réfléchie du *Big Data* et de l'apprentissage automatique peut apporter à la clinique et à la recherche.

3. Prédire

L'étude précédente introduit l'aspect du *Big Data* qui suscite peut-être le plus d'espoir, mérité ou non : la capacité non seulement de détecter et donc de classer, mais celle aussi de donner une prédiction individuelle. Ce défi clinique, de connaître à l'avance le cours, l'évolution de la maladie pour *un patient donné*, et ainsi de savoir ce qui aidera ce patient en particulier, concerne à différents degrés toute la médecine. Il intéresse également le politique et l'économie, qui l'érigent comme principe au coeur des efforts de transformation des systèmes de soins (Présidence de la République, 2018b). Il ferait peut-être aussi écho aux attentes d'une société craintive face à l'incertitude, moins tolérante face aux aléas rappelant la fragilité de nos mécanismes de contrôle. La notion de prédiction se décline en plusieurs volets : la prévention primaire et secondaire, et l'optimisation du traitement ou du suivi par la prédiction de la réponse individuelle au traitement et la personnalisation du choix du traitement. En psychiatrie, cela renvoie à des préoccupations majeures autour de la pratique clinique : prévention du passage à l'acte

suicidaire ou autre ; prévention d'une décompensation clinique chez un patient psychotique ou bipolaire ; identification parmi des personnes à haut-risque, de celles qui évolueront vers un trouble autistique, une schizophrénie ou une dépendance à une substance psychoactive, afin de mettre en place des interventions précoces ; prédiction de la réponse au traitement anti-dépresseur ou anti-psychotique afin d'orienter le choix thérapeutique.

a. Prévention primaire et secondaire : prédire la survenue et le cours de la maladie

Comme nous l'avons vu (voir ci-dessus II.2.e., p. 100), l'étude de Whelan et al., en analysant une variété de données, retrouve une contribution multifactorielle à la prédiction du mésusage d'alcool chez les adolescents, suggérant une intrication étroite – voire une superposition ? – entre une histoire biologique, une histoire narrative de vie, et une subjectivité en développement (Whelan, Watts, Orr, Althoff, Artiges, Banaschewski, Barker, Bokde, et al., 2014). Ceci laisse en suspens la question étiologique, les facteurs décrits n'étant peut-être que des stigmates d'un processus étiologique non-encore identifié. Mais cet outil a le mérite d'orienter l'attention vers des pistes d'étude, et de fournir des points de départ pour la conception d'éventuelles interventions préventives face à une pathologie aux conséquences personnelles et sociales lourdes.

Dans le même esprit, le poids de la morbidité globale due au trouble dépressif majeur (Vos et al., 2017) motive la recherche de stratégies de prise en charge plus efficaces. L'une de ces pistes – mieux prédire le cours de la maladie pour prévenir une évolution chronique ou une rechute – est explorée par certaines études utilisant des outils d'apprentissage statistique avec des données multimodales. Les études réellement prédictives, c'est-à-dire prospectives, avec un suivi longitudinal, sont moins nombreuses que les études de classification, et pour le trouble dépressif, elles recherchent davantage la réponse au traitement plutôt que l'histoire naturelle de la maladie (voir ci-dessous, sous-chapitre II.3.b. p. 112). Cependant, on peut citer comme exemples les études de Schmaal et al. et Farb et al., toutes deux néanmoins sans validation externe et avec des cohortes de faible taille (Farb, Anderson, Bloch, & Segal, 2011; Schmaal et al., 2015). Schmaal et al. suivent 118 patients dépressifs (issus de la cohorte néerlandaise NESDA pour l'étude de la dépression) pour une durée de deux ans, après le recueil de données cliniques et

d'IRM morphologiques et fonctionnelles (fonctions exécutives et réponse aux expressions faciales des émotions). Utilisant des *Gaussian process classifiers*, ils recherchent les facteurs prédictifs pour trois trajectoires différentes de la maladie, précédemment définies à partir d'une cohorte plus élargie par *latent-class growth analysis*. Distinguant 23 patients avec une évolution chronique, 36 avec une amélioration progressive et 59 avec une rémission rapide, ils obtiennent la meilleure discrimination entre patients chroniques et patients en rémission rapide (précision de 73 %), avec les données de la tâche émotionnelle en IRMf. Cette même tâche distingue également, mais faiblement, l'évolution chronique de l'amélioration progressive. Les autres modalités (clinique, IRMf en fonctions exécutives, IRM morphologique) ne permettent pas une distinction suffisante entre les différentes trajectoires, ce qui n'étonne guère dans le cas des données cliniques utilisées, qui sont assez grossières (sévérité et durée de l'épisode dépressif, comorbidités, résultat de score de personnalité, présence de traumatisme dans l'enfance). De même, la combinaison des données cliniques avec les données d'imagerie n'augmente pas le pouvoir prédictif, indiquant que la multi-modalité n'a d'intérêt que si le « signal » est supérieur au « bruit » généré par l'excès de données.

La cohorte étudiée par Farb et al. est encore plus restreinte, composée de 16 patients dépressifs en rémission et 16 sujets sains, suivis pendant 18 mois après le recueil de données d'IRMf (visionnage d'une vidéo à valence émotionnelle), de questionnaires sur le fonctionnement cognitif/affectif et d'échelles sémiologiques (Farb et al., 2011). Dix patients parmi les 16 rechutent au cours de la période de suivi. Par régression logistique sur les données d'imagerie, les auteurs identifient la réactivité plus importante du cortex préfrontal médial comme prédictif de rechute, tandis que l'activation du cortex visuel semble protecteur. La normalisation de l'activation du cortex préfrontal au niveau observé chez les sujets sains est liée à l'absence de rechute. La réactivité du cortex préfrontal médial serait associée à une tendance à la rumination (mesurée par un questionnaire cognitivo-affectif), alors que l'activation prédominante du cortex visuel serait liée à une tendance à mieux accepter les situations de détresse, suggérant peut-être les différences cognitives et affectives sous-tendant la divergence des trajectoires pathologiques. Mais – on ne se lassera pas de le répéter – la valeur prédictive de ces résultats ne pourra être établie que par réplication sur une cohorte de validation externe.

D'incidence inférieure au trouble dépressif, mais peut-être plus marquant de par sa violence, le suicide est une cible évidente des études de prédiction, car malgré la quantité de connaissances épidémiologiques, leur mise en opération pour cibler les efforts de

prévention, ou pour l'évaluation du risque individuel, reste difficile (McCoy, Castro, Roberson, Snapper, & Perlis, 2016; Olfson, Marcus, & Bridge, 2014). Niculescu et al. proposent un outil prédictif composite, comprenant d'une part deux échelles cliniques simples accessibles sous la forme d'applications numériques (une échelle affective et une échelle recueillant les facteurs démographiques, anamnestiques, autobiographiques, culturels influençant le risque suicidaire), d'autre part onze biomarqueurs d'expression génique associés au suicide (Niculescu et al., 2015). Au sein d'une cohorte longitudinale de 217 hommes ayant des troubles psychiatriques variés, 37 personnes constituent la cohorte de découverte, suivie à intervalles réguliers avec un recueil à chaque rencontre, de données cliniques, de la présence d'idées suicidaires, et d'un prélèvement sanguin pour l'analyse transcriptomique. Les auteurs espèrent ainsi suivre conjointement les variations d'expression génique et l'idéation suicidaire de manière longitudinale chez chaque sujet individuel. Les biomarqueurs identifiés dans cette cohorte sont ensuite validés dans une cohorte post-mortem de 26 suicidés. Puis la valeur prédictive de l'outil composite est testée dans deux cohortes différentes, pour prédire l'apparition d'idées suicidaires (cohorte de 108 patients) et l'hospitalisation pour idées suicidaires (cohorte de 157 patients).

L'approche ici diffère de celles vues tout au long de ce chapitre, car au lieu de rechercher les facteurs prédictifs et les biomarqueurs par apprentissage automatique, les auteurs choisissent de les identifier « manuellement », en parcourant la littérature et les bases de données génomiques pour trouver des gènes dont les variations d'expression sont associées au suicide, puis en établissant un système manuel de pondération pour attribuer un score plus ou moins élevé à chaque gène en fonction de la fiabilité de son association au suicide, et de la cohérence des données individuelles observées dans la cohorte de découverte. De même, les échelles cliniques ont été précédemment construites et décrites par les auteurs. Les biomarqueurs identifiés sont impliqués dans une variété de processus cellulaires et physiologiques, de voies de signalisation et de pathologies (inflammation, réponse immunitaire, apoptose, connectivité neuronale, humeur, psychose etc). L'outil composite incluant les modalités cliniques et biologiques, quant à lui, semble avoir une excellente performance pour prédire la survenue d'idées suicidaires dans la cohorte de validation en question (AUC 0,92) et une performance moyenne pour prédire une hospitalisation (AUC 0,71). Les auteurs expliquent leurs méthodes de manière détaillée, donnant des justifications soutenant leurs choix et indiquant certaines de leurs limites. Le choix d'une pondération manuelle peut être

critiquée, tout comme elle peut être une force, par la transparence du processus de découverte. La petite taille des cohortes est une faiblesse, mais la diversité des cohortes de validation et l'inclusion d'une validation post-mortem sont des atouts importants dans cette étude. Sans l'utilisation d'outils d'apprentissage automatique, les auteurs ont pu appréhender plusieurs cohortes hétérogènes, des sources de données complexes (haute-dimensionnalité du transcriptome, croisement avec le facteur temps dans les suivis longitudinaux individuels), et multi-modales, montrant qu'une approche athéorique n'est pas la seule possible. En pratique, les questionnaires/échelles de l'outil composite seraient plus utiles et facilement applicables en clinique (et ont une bonne efficacité de prédiction à eux seuls, même si ce résultat n'est pas le critère de jugement principal de l'étude). Le versant biologique contribuerait davantage à des efforts de recherche. En effet, il semble difficile d'envisager actuellement d'effectuer une analyse d'expression génique chez tout patient psychiatrique ou tout suicidant pour juger du risque de passage à l'acte.

Revenant aux outils d'apprentissage automatique, et aux bases massives tant en dimensions qu'en effectifs, Kessler et al. établissent un modèle de prédiction des décès par suicide dans les suites d'une hospitalisation pour motif psychiatrique, chez les soldats de l'armée américaine (Kessler et al., 2015). Souhaitant répondre à l'augmentation de l'incidence de suicide observée depuis 2004 parmi les soldats, ils utilisent certaines bases administratives militaires, comprenant un nombre important de données démographiques, sociales, professionnelles (carrière militaire), judiciaires et médicales. Sur une période de cinq ans, ils retrouvent plus de 50 000 hospitalisations psychiatriques, concernant plus de 40 000 soldats, et 68 décès par suicide dans les douze mois suivant une hospitalisation. L'identification des variables prédictives se fait par arbres décisionnels multiples en régression (*multiple regression trees*), révélant 421 variables multi-colinéaires qui sont ensuite raffinées par *elastic net* et *forward stepwise selection* pour diminuer la variance du modèle. Des modèles de survie logistique multivariés sont estimés à partir de ces variables : le modèle le plus efficace et stable utilise 20 variables (AUC 0,84), et les plus importants prédicteurs de décès par suicide dans les douze mois après hospitalisation psychiatrique appartiennent à la fois aux catégories médicales et sociales. Le sexe masculin, les antécédents judiciaires, la possession d'une arme à feu, l'âge au recrutement supérieur à 27 ans, ainsi que les antécédents d'idées ou de tentatives de suicide, les antécédents de suivi, de diagnostic et d'hospitalisation en psychiatrie (notamment dépression et anti-dépresseurs) et la cause de la dernière hospitalisation (psychose et troubles somatoforme et conversif, notamment), sont les variables les plus

fortement prédictives. Distinguant quatre niveaux de risque, les auteurs indiquent que le niveau de risque le plus élevé comprend les 5 % d'hospitalisations qui correspondent à plus de 50 % des suicides.

Focalisant sur une population à risque (soldats post-hospitalisation psychiatrique), avec une base de données conséquente et interconnectée, mais pas toujours fiable (base administrative et non de recherche, données manquantes ou erronées), cette étude montre l'intérêt du croisement des données, mais aussi la difficulté d'étudier le suicide. Avec un évènement aussi statistiquement rare (68 évènements après cinq ans de recueil de données concernant plus de 50 000 hospitalisations), un outil prédictif risque d'avoir de nombreux faux positifs. Et inversement, la sélection d'une population déjà à risque fait l'impasse sur la proportion importante de suicides survenant chez les personnes jamais hospitalisées, ni soignées en psychiatrie. En effet, les auteurs notent que les 68 suicides correspondent à 12 % des suicides enregistrés chez les soldats durant la période d'étude. Malgré la concordance de certains prédicteurs cités ici avec des facteurs de risque identifiés par des études dans des populations civiles (Schoenbaum et al., 2014), la généralisation de ce modèle précis à la population générale est discutable, d'autant plus que la bureaucratie civile demeure – pour l'instant – moins englobante et moins aisément interconnectée que la bureaucratie militaire.

À ce titre, les études de McCoy et al. et Simon et al., utilisant les dossiers de soins informatisés comme sources de données croisées avec les registres de décès, offriraient des prédictions plus facilement applicables grâce à la disponibilité de ces modalités de données et leur automatisation possible, et plus généralisables de par la meilleure représentativité des cohortes étudiées (McCoy et al., 2016; Simon et al., 2018). Dans une étude rétrospective examinant neuf ans de données cliniques concernant tous les patients ayant été hospitalisés dans deux hôpitaux universitaires américains, McCoy et al. utilisent à la fois les données administratives et de codage, ainsi que le texte libre des comptes rendus d'hospitalisation analysé par NLP pour y repérer une valence émotionnelle négative ou positive. Cet outil de traitement du langage naturel avait été précédemment utilisé par ces auteurs pour prédire la ré-hospitalisation et la mortalité suite à une hospitalisation (McCoy et al., 2015). En croisant les données hospitalières avec les registres publics de décès, les auteurs retrouvent 235 suicides, parmi plus de 800 000 hospitalisations impliquant plus de 400 000 patients. Le modèle de survie (plusieurs sont testés par les auteurs, notamment *Cox regression*) identifie, à partir des données administratives, plusieurs variables d'intérêt, socio-démographiques (race blanche, sexe

masculin) et médicales (nombre élevé de passage aux urgences et de consultations psychiatriques). La valence émotionnelle positive des compte-rendus de sortie est associée à une diminution du risque de suicide, alors que la valence négative l'augmente, mais l'inclusion de l'analyse NLP n'améliore que peu le modèle obtenu à partir des données administratives.

Au sein de sept groupements de santé dans différents états américains, Simon et al. tentent de prédire le risque de suicide et de tentative de suicide à 90 jours, chez près de 3 millions de patients ayant eu une visite ambulatoire en soins psychiatriques (plus de 10 millions de visites) ou en soins généralistes avec un diagnostic psychiatrique (plus de 9 millions), sur une période de six ans (Simon et al., 2018). Ils identifient sur cette période, plus de 24 000 tentatives de suicide et 1240 décès par suicide, dans les 90 jours après une visite. Se basant sur les données socio-démographiques et médicales codées dans les dossiers médicaux, incluant quand disponible le questionnaire PHQ-9 (*Patient Health Questionnaire*), des modèles de régression logistique, avec sélection de variables par la technique lasso, sont développés sur 65 % de la cohorte choisie au hasard, laissant les 35 % restants comme cohorte de validation indépendante. Sans surprise, les variables les plus prédictives de suicide et de tentative sont les antécédents de tentative de suicide, de diagnostic psychiatrique ou addictologique, d'hospitalisation ou de recours aux urgences psychiatriques et la réponse à l'item 9 du PHQ (sur les idées suicidaires). On retrouve ici aussi que les 5 % de visites au risque le plus élevé correspondent à plus de 40 % des suicides et des tentatives, à la fois suite aux visites spécialisées et aux visites généralistes. Les modèles prédictifs montrent une bonne stabilité entre les cohortes d'entraînement et de validation, et dans les deux cas, une bonne performance (statistique-C supérieure à 0,8) qui, selon les auteurs, dépasse celle d'autres outils disponibles, comme l'item 9 du PHQ-9 par exemple. En parallèle avec la faible incidence du suicide (parmi les 5 % de patients les plus à risque, 5,4 % tentent de se suicider et 0,26 % décèdent par suicide), la valeur prédictive positive est faible, atteignant un maximum de 10 %. Les résultats de cette étude sont intéressants, non parce qu'ils sont une révélation – au contraire, les facteurs prédictifs retrouvés ici sont des facteurs de risque bien connus – mais par la démonstration qu'un outil prédictif avec une bonne performance peut être construit à partir de données en soi très simples, par leur mise en relation complexe au sein d'un échantillon très large. L'étude a des faiblesses, comme le choix de cohorte de validation (la mise à l'écart d'un site spécifique aurait donné plus d'indépendance à la cohorte), le choix de patients ayant un diagnostic psychiatrique, et d'autres considérations

méthodologiques autour de l'identification des variables prédictives et la présentation des résultats. Mais, comme l'indiquent les auteurs, les modèles développés ainsi que d'autres données sont librement mis à disposition afin de faciliter la réplication dans d'autres cohortes et d'autres conditions.

L'espoir d'amélioration du pronostic par des interventions précoces bien ciblées stimule la recherche de meilleurs outils prédictifs pour les troubles autistiques et psychotiques, ce d'autant plus que leur apparition souvent progressive avec l'existence d'une période prodromique, et le jeune âge des personnes touchées laissent croire à la possibilité de prévention primaire. La non-spécificité du tableau clinique précoce d'autisme par rapport à d'autres troubles de la petite enfance (Glatt et al., 2012; Pramparo et al., 2015), et la valeur prédictive positive faible des évaluations cliniques dans les états à risque de psychose (Schmidt et al., 2017), montrent les limites des outils existants. De ce fait, pour les troubles du spectre autistique, la recherche s'oriente vers des prédicteurs biologiques, comme l'imagerie (Hazlett et al., 2017) et l'expression génique (Glatt et al., 2012; Pramparo et al., 2015). L'étude de Hazlett et al., évoquée en début de chapitre (voir sous-chapitre II.1, p. 68), identifie des trajectoires de développement cérébral différentes chez les enfants à haut-risque familial qui seront ensuite diagnostiqués autistes, par rapport à leurs pairs à haut et bas-risque non-autistes. Un algorithme d'apprentissage profond sur les données d'IRM morphologique à six et douze mois permet, de surcroît, de prédire un diagnostic positif à l'âge de deux ans chez 30 enfants parmi 34 dans une cohorte de 179 enfants à haut-risque, donnant une valeur prédictive positive de 81 %. Le modèle n'est pas validé sur une cohorte externe, mais il est librement disponible pour être utilisé dans une étude de réplication. Compte tenu de la rapidité du développement cérébral et psychique chez les enfants à cet âge (six à 24 mois), une des qualités de cette étude est le suivi prospectif, avec des instances multiples d'observations complètes (clinique et imagerie), permettant une vision longitudinale réellement individualisée, plutôt qu'une moyenne sur des enfants différents.

Le modèle n'est vraisemblablement pas généralisable à tous les enfants présentant un retard de développement, tant sur le plan de la validité statistique, que sur le plan pratique. L'étude s'est focalisée sur une population à haut-risque, à juste titre car on imagine mal administrer une IRM à tous les enfants présentant un retard de langage, notamment à un âge déjà trop avancé pour capter les modifications développementales précoces qui contribuent à ce modèle prédictif. En revanche, au sein d'une famille avec un enfant autiste, un tel modèle pourrait avoir du sens.

Les études d'expression génique viennent pallier à ces défauts (Glatt et al., 2012; Pramparo et al., 2015) : les cohortes sont plus représentatives des populations rencontrées en pratique clinique, et la modalité d'examen, un prélèvement sanguin, peut être déployée plus facilement chez des enfants en bas-âge. Il s'agit d'études réalisées par la même équipe de recherche, l'étude de 2012 étant une étude pilote. Dans les deux cas, il s'agit d'un recrutement au sein de la population générale, par un outil de dépistage pour les troubles du développement (questionnaire CSBS DP, *Communication and Symbolic Behavior Scales Developmental Profile*) adopté par des pédiatres de la communauté. Alors que Glatt et al. ont une cohorte mixte, Pramparo et al. ne sélectionnent que les garçons pour diminuer l'hétérogénéité de la cohorte. Les deux études sont prospectives, recrutant des enfants âgés de un à quatre ans, avec un prélèvement sanguin au recrutement et un suivi régulier jusqu'à la confirmation du diagnostic. L'analyse du transcriptome leucocytaire recherche les différences d'expression génique entre les enfants qui auront un diagnostic d'autisme, et ceux qui auront un développement normal ou une anomalie non-autistique (retard de langage simple, retard global de développement), le recrutement communautaire prospectif permettant justement de capter cette population correspondant aux diagnostics différentiels. Après identification et sélection/réduction de variables, Glatt et al. construisent un modèle SVM à partir des marqueurs géniques exprimés différemment, et Pramparo et al. construisent un modèle par régression logistique sur des modules de co-expression génique (identifiés par clustering).

Glatt et al. ont une cohorte de 179 enfants (60 autistes, 34 avec retard de langage simple, 17 avec retard global de développement, 68 développement normal), qu'ils choisissent de diviser au hasard pour constituer une cohorte d'entraînement et une cohorte de validation. Pramparo et al. recrutent 142 enfants pour leur cohorte d'entraînement (87 auront un diagnostic d'autisme, 41 avec un développement normal, 14 avec un trouble du développement autre), et 73 enfants provenant de la cohorte de Glatt et al. sont retenus comme cohorte de validation, créant ainsi une certaine indépendance par rapport à la cohorte d'entraînement (44 autistes, 25 à développement normal, 4 avec trouble autre). Les deux études identifient, entre autres, des gènes impliqués dans la réponse immunitaire. Le modèle de Glatt et al. obtient une précision de 87 % pour la prédiction de l'autisme dans la cohorte de validation. Le modèle de Pramparo et al. atteint une précision de 75 % (spécificité 72 %, sensibilité 77 %) pour prédire un diagnostic

d'autisme dans la cohorte de validation, et semble spécifique à ce trouble car il ne prédit pas efficacement la survenue des autres troubles du développement.

Les auteurs argumentent qu'en comparaison aux outils de dépistage existants, comme le questionnaire M-CHAT (*Modified Checklist for Autism in Toddlers*) dont la spécificité est faible dans la population générale (27 %), leurs outils de prédiction sont une avancée importante. Mais face à deux études avec autant de similitudes, réalisées par la même équipe, et dont l'une est de surcroît une étude pilote, on ne peut s'empêcher de penser que l'intérêt scientifique et clinique aurait été mieux servi par une approche plus réfléchie : la validation du premier modèle dans la nouvelle cohorte, la construction du nouveau modèle sur l'ancienne cohorte, ou la comparaison qualitative et quantitative précise des deux modèles avec d'autres mesures de performance, voire même une synthèse des deux modèles. Ceci aurait permis de confronter les méthodologies, de mieux tester la fiabilité des modèles et des prédicteurs (comprendre si les variations sont dues au choix des outils, aux variations entre cohortes etc), et donc de mieux cerner leur intérêt pratique et scientifique. Cette fragmentation et l'impression d'un travail sisyphéen rappellent la situation des études de classification en imagerie, telle que la critiquent Woo et al. (voir sous-chapitre II.2.b., p.85).

Dans ce contexte, rediriger certains efforts vers des travaux de synthèse et de méta-analyse, comme celle de Schmidt et al. concernant le pronostic des états à haut-risque de psychose, semble indispensable (Schmidt et al., 2017). La recherche sur les marqueurs du prodrome psychotique inclut une diversité de sources et de modalités : cliniques (Cannon et al., 2016; Carrión et al., 2016; Ciarleglio et al., 2018), socio-démographiques (Cannon et al., 2016; Carrión et al., 2016; Nieman et al., 2014), cognitives (Cannon et al., 2016; Carrión et al., 2016; Koutsouleris, Davatzikos, et al., 2012; Pettersson-Yeo et al., 2013), électrophysiologiques (Nieman et al., 2014), biochimiques/sériques (Chan et al., 2015; Perkins et al., 2015), linguistiques (Bedi et al., 2015), imagerie morphologique, fonctionnelle et de diffusion (Koutsouleris, Borgwardt, et al., 2012; Pettersson-Yeo et al., 2013). Dans leur revue systématique des travaux publiés jusqu'à Octobre 2015 sur la prédiction du pronostic dans les états à haut-risque de psychose, Schmidt et al. identifient 25 travaux, utilisant des données cliniques (n=10), biologiques (imagerie, EEG, protéines sériques, n=5), neurocognitives (n=5), environnementales (n=5), dont certaines combinent différentes modalités (n=8). Les études présentent des cohortes de taille variable, allant de 30 à plus de 300 sujets, avec des variations également dans l'âge des sujets et la durée de suivi. Les outils statistiques

utilisés sont divers, mais il y a une majorité nette de modèles par régression logistique et régression de Cox. Aucune des études ne présente une validation externe, certaines ayant néanmoins une étape de validation interne, mais d'autres n'ayant aucune validation.

Les modèles les plus performants dans les domaines cliniques, biologiques, neurocognitifs, y compris parmi les études combinant différentes modalités, atteignent des valeurs prédictives positives supérieures à 82 %. Les « meilleurs » modèles sont basés sur une combinaison de certains symptômes positifs et négatifs en clinique, les volumes de substance grise en imagerie, les fonctions verbales et exécutives en neurocognition, et une combinaison de symptômes de dissociation psychique, de fonctionnement verbal et social dans le modèle incluant différentes modalités. Les modèles contenant uniquement des données environnementales présentent une performance nettement moindre, ne dépassant pas 63 % en valeur prédictive positive, qui peut s'expliquer par le manque de spécificité et de granularité des données utilisées.

À la recherche d'une synthèse pratique de ces résultats, Schmidt et al. essaient d'estimer si une application séquentielle de ces différents modèles permettrait de mieux prédire quels sujets à haut-risque auront bien une transition psychotique, améliorant la spécificité du diagnostic d'état à haut-risque. Ils font une simulation statistique à partir des modèles avec validation interne, pour déterminer quelle combinaison de trois tests donnerait la meilleure valeur prédictive positive pour prédire la survenue de psychose. Il s'agit d'une simulation purement statistique, probabiliste (application du théorème de Bayes sur les caractéristiques de sensibilité, spécificité, valeurs prédictives positives et négatives des différents modèles considérés), et non pas un calcul basé sur une application réelle de la séquence de tests sur une cohorte. La meilleure séquence met en jeu d'abord un modèle combinant des données de potentiels évoqués et des données cliniques (Nieman et al., 2014), puis un modèle d'imagerie morphologique (Koutsouleris et al., 2015), et enfin un modèle basé sur la détection de protéines sériques (Perkins et al., 2015). Si les trois tests sont positifs, la valeur prédictive positive obtenue est de 98 %, donnant une quasi-certitude quant au risque de transition psychotique. Bien entendu, ce résultat provient d'une simulation statistique, et non de données expérimentales directes. Les modèles sur lesquels il est basé nécessitent tous une validation externe, donc les valeurs prédictives positives obtenues sont probablement surestimées. Mais cette méta-analyse offre une stratégie concrète d'exploration diagnostique et pronostique, et remet en contexte des modèles disparates, tissant quelques liens entre les manifestations cliniques et les modifications biologiques.

Cette méta-analyse n'a pas pu prendre en compte un autre modèle important, publié après 2015, ayant fait l'objet d'une validation externe publiée en parallèle (Carrión et al., 2016). Issu de la North American Prodrome Longitudinal Study (NAPLS), ce modèle utilise des variables sélectionnées a priori et manuellement sur la base de leur publication antérieure dans la littérature, et non par apprentissage automatique (Cannon et al., 2016). Le but des auteurs étant de créer un outil prédictif facilement généralisable en pratique clinique courante, et utilisable dès le premier contact avec le patient, ils se sont limités à la recherche de variables cliniques, démographiques, cognitives et de fonctionnement global, identifiant des échelles et des questionnaires pertinents dans la littérature. Le nombre de variables est restreint a priori à huit, afin de maintenir un nombre de variables nettement inférieur au nombre de sujets, pour éviter l'état de haute-dimension et le risque de surapprentissage.

Dans les deux ans de suivi de 596 sujets à haut-risque, 84 personnes font la transition vers un diagnostic de psychose (16 % de la cohorte). Le modèle, une analyse de survie par régression de Cox multivariée, retrouve comme facteurs prédictifs des symptômes positifs atténués, une baisse du fonctionnement social, des perturbations du fonctionnement cognitif et un âge plus jeune. Les antécédents familiaux de psychose, ou personnels de traumatisme ou facteurs de stress ne sont pas des facteurs prédictifs significatifs dans ce modèle. Plutôt qu'une classification catégorique, le modèle fournit une probabilité individuelle de transition psychotique. Ainsi, la valeur prédictive positive est maximale (48 %) si le seuil de probabilité choisi pour le pronostic de transition est à 0,4. Mais à ce niveau de risque, la sensibilité est faible (seuls 17 % des cas de transition atteignent cet niveau de risque). Au seuil de risque de 0,2, la sensibilité augmente à 66 %, capturant les deux tiers des cas de transition, mais avec de nombreux faux positifs (valeur prédictive positive à 28 %).

Toujours dans l'optique de faciliter son utilisation, les auteurs mettent à disposition ce modèle sous la forme d'un calculateur accessible par internet. Ce résultat est soutenu par la taille de la cohorte, la publication concomitante d'une validation externe, et les efforts pour limiter le risque de surapprentissage. La sélection manuelle des variables et la limitation de leur nombre offre une certaine transparence et une intelligibilité au modèle, que l'on ne retrouve pas forcément dans d'autres modèles. Le risque, cependant, est celui d'un biais trop important. Et l'on voit, en effet, que la prédiction positive n'est pas encore assez performante pour permettre, par exemple, à un clinicien de trancher la question d'un

traitement. Les auteurs reconnaissent ceci en recommandant, pour l'instant, l'utilisation de cet outil à des fins de recherche, dans des essais thérapeutiques notamment.

b. Optimisation du traitement : prédire la réponse à une intervention

Compte tenu du délai nécessaire pour pouvoir juger de l'efficacité d'une intervention thérapeutique, médicamenteuse ou autre, en psychiatrie (plusieurs semaines au moins), le choix d'un traitement qui s'avère inefficace peut être très délétère, laissant le patient aux prises d'un processus pathologique, tout en l'exposant à d'éventuels effets indésirables, sans effet bénéfique. La question de l'optimisation du traitement a, sans conteste, un intérêt économique sous-entendu dans ce libellé, mais il s'agit avant tout d'une question humaine et déontologique, très pertinente en psychiatrie où les taux de non-réponse aux traitements de première intention peuvent être élevés (Gillan & Whelan, 2017). La recherche de facteurs intrinsèques et extrinsèques influençant la réponse aux traitements a jusque-là, essentiellement pris la forme d'études de corrélations au niveau de cohortes entières, plutôt que de prédictions individuelles, comme le recensent la majorité des études dans la revue de Stern et al. sur la réponse aux traitements médicamenteux dans le trouble bipolaire, la schizophrénie, la dépression et l'autisme (Stern et al., 2018).

Parmi les études de cette revue utilisant des modélisations prédictives, la majorité sont préliminaires, avec des cohortes de petite taille, des groupes témoins inadapés, et ne montrant pas d'étape de validation interne du modèle, sans parler d'une validation externe. Les variables examinées appartiennent à la clinique au sens large (y compris des échelles neurocognitives), la génétique, l'imagerie, et parfois à d'autres sources comme l'électrophysiologie. Il est difficile de tirer davantage d'informations de la plupart de ces études, sans tomber dans la spéculation, en raison de leur qualité variable et des critères de validité manquants. Deux résultats intrigants ressortent néanmoins de deux études génétiques : l'un issu d'une étude GWAS par un consortium international sur la pharmacogénétique du lithium, retrouve une moindre efficacité du lithium chez les patients bipolaires ayant un score polygénique élevé pour la *schizophrénie*, remettant en relief la pléiotropie à l'évidence dans la génétique de ces deux pathologies, et suggérant peut-être une spécificité du lithium pour le versant purement thymique du trouble bipolaire (International Consortium on Lithium Genetics (ConLi+Gen) et al., 2018). L'autre résultat,

négatif cette fois, ne retrouve aucun marqueur génétique individuel ou score polygénique significatifs, prédictifs de la réponse aux antidépresseurs, dans deux cohortes différentes à ascendance européenne (Tansey et al., 2012). Si tant est qu'il y ait des prédicteurs pharmacogénétiques fiables dans la dépression, la taille des cohortes, atteignant pourtant ici plusieurs milliers de sujets, est tout de même peut-être insuffisante pour détecter un effet significatif, dans un trouble aussi cliniquement hétérogène, sans hypothèse a priori ciblant les recherches.

Les traitements médicamenteux ne sont qu'une des modalités thérapeutiques en psychiatrie. Plusieurs études cherchent, par exemple, à prédire l'efficacité de psychothérapies à partir de marqueurs en imagerie. Une étude préliminaire de 38 patients, sans validation externe, identifie des schémas de connectivité fonctionnelle et anatomique (IRMf de repos et IRM de diffusion) liés à l'efficacité de la thérapie cognitivo-comportementale (TCC) dans l'anxiété sociale, les différentes modalités d'imagerie améliorant nettement un prédicteur basé sur l'évaluation clinique seule (Whitfield-Gabrieli et al., 2016). Une autre étude recueille les IRMf de 49 patients dépressifs, divisés en deux petites cohortes, au cours d'une tâche émotionnelle, avant de débiter un traitement par thérapie cognitive seule, puis une nouvelle fois après le traitement (Siegle et al., 2012). Une analyse par forêt d'arbres décisionnels identifie l'activité pré-thérapeutique du cortex cingulaire antérieur et du cortex pré-frontal dorsolatéral, ainsi que la sévérité initiale comme prédicteurs possibles, les patients ayant une réactivité du cortex cingulaire plus faible au départ montrant une meilleure réponse à la thérapie cognitive. Le cortex cingulaire est aussi impliqué dans une étude évaluant la réponse au traitement par électroconvulsivothérapie (ECT) lors d'un essai non-randomisé de 67 sujets (46 patients dépressifs et 21 sujets sains) avec utilisation de données d'imagerie morphologique pour construire des modèles SVM (Redlich et al., 2016). Un volume moindre du cortex cingulaire sur l'IRM pré-thérapeutique prédirait une moindre efficacité du traitement par ECT dans ce groupe de patients. Pouvoir distinguer, parmi plusieurs classes de médicaments (ou de modalités thérapeutiques), celle qui sera bénéfique au patient, est d'une plus grande utilité, car cela répond davantage à la problématique telle qu'elle se pose en pratique. Deux essais tentent d'aborder la question, utilisant des données d'évaluations cognitives (Etkin et al., 2015) et d'EEG (DeBattista et al., 2011). Dans une cohorte de 665 patients dépressifs randomisés sur différents traitements médicamenteux, Etkin et al. identifient un sous-groupe au fonctionnement cognitif altéré dans la plupart des treize tests recueillis avant traitement par les auteurs. Au sein de ce groupe, les mesures

cognitives semblent contribuer au pronostic de rémission sous escitalopram, sans pouvoir prédire la réponse aux autres traitements. DeBattista et al. utilisent le système de *referenced-EEG* (Suffin et al., 2007), qui fonde ses recommandations de traitement antidépresseur sur des mesures quantitatives d'EEG, en se reposant sur une base d'environ 1800 EEG individuels mis en corrélation avec des résultats de traitement médicamenteux (DeBattista et al., 2011). Dans un essai randomisé multicentrique incluant 89 patients avec une dépression résistante, les auteurs trouvent que les traitements recommandés par ce système sont plus efficaces que ceux recommandés dans un très large essai clinique sur la dépression résistante.

Ces études montrent une diversité de modalités de traitement et de marqueurs. Cependant, aucun n'a jusqu'ici abouti à une utilisation répandue en clinique, peut-être justement parce qu'il y a trop de diversité (dans les questions posées, les interventions, le choix du groupe témoin etc). La rigueur méthodologique nécessaire pour trouver un prédicteur efficace, généralisable, facilement applicable à large échelle et répondant pertinemment à une problématique clinique, est telle qu'on la retrouve dans peu d'études pour l'instant. À cet égard, Gillan et Whelan proposent un certain nombre de principes directeurs pour améliorer la recherche prédictive (Gillan & Whelan, 2017), principes qui font écho aux faiblesses que l'on a pu voir dans les études évaluées tout au long de cette revue, et qui rappellent certaines des propositions de Woo et al. en imagerie (Woo et al., 2017) :

- taille des cohortes : le risque de sur-apprentissage (et donc d'un modèle non-généralisable) est important dans les petites cohortes, d'où la nécessité d'utiliser de larges cohortes, en ayant recours à des consortiums d'étude par exemple.
- étape cruciale de la validation : en plus d'une validation interne incontournable, la validation externe est indispensable lorsque plusieurs techniques d'apprentissage automatique sont essayées pour trouver un modèle performant, la validation externe sur une cohorte indépendante étant encore plus fiable.
- mesures de performance : entre AUC, sensibilité, spécificité, valeurs prédictives positives et négatives, rapports de vraisemblance positifs et négatifs, une seule mesure n'est souvent pas suffisante pour récapituler la performance d'un modèle. Les mesures rapportées doivent, notamment, pouvoir répondre de manière pertinente à la pratique clinique visée par le prédicteur (dépistage, diagnostic positif, orientation thérapeutique etc).

- termes de « regret » : pour mieux refléter la réalité que doit appréhender le jugement clinique, les modèles pourraient incorporer des pénalités sur les erreurs de prédiction qui mènent à des conséquences sévères (comme un faux positif menant à une intervention chirurgicale lourde).
- transparence/interprétabilité/intelligibilité : reprenant les recommandations citées précédemment (Woo et al., 2017), les facteurs prédictifs devraient être résumés de manière humainement intelligible, et évalués sur leur degré de cohérence avec les corpus de connaissances en neurosciences, et sur l'existence de facteurs confondants.
- Open Science : le partage des données, et plus encore, des modèles, est un élément essentiel des efforts de validation et de réplication.

On rappellera ici la rigueur méthodologique de l'étude de classification et de prédiction effectuée par Drysdale et al. auprès de patients dépressifs résistants, identifiant des biotypes de dépression correspondant à des schémas distincts de connectivité fonctionnelle, et à des réponses différentes au traitement par rTMS (Drysdale et al., 2017). L'intégration du biotype et des données de connectivité fonctionnelle permet une meilleure prédiction de la réponse au traitement par rTMS (ciblant les aires pré-frontales dorso-médiales), que la prédiction sur les données cliniques seules (plus de 85 % de précision, y compris sur une cohorte indépendante, contre 62 % de précision pour les données cliniques).

Chekroud et al. et Koutsouleris et al. sont deux autres études de prédiction montrant une méthodologie rigoureuse (Adam Mourad Chekroud et al., 2016; Koutsouleris et al., 2016). L'élaboration de prédicteurs d'efficacité des antidépresseurs, à partir de données cliniques provenant des cohortes de larges essais cliniques multicentriques, est un approche préalablement utilisée par Uher et al., qui utilisent, pour ce faire, les données individuelles de deux essais importants d'antidépresseurs, un essai pour l'entraînement du modèle prédictif (GENDEP), le deuxième (STAR*D) pour la validation (Uher et al., 2012). Ils avaient auparavant représenté la symptomatologie dépressive par neuf dimensions de symptômes identifiées à partir d'une analyse factorielle d'échelles cliniques de dépression. Ils modélisent ainsi l'influence de ces facteurs sur la réponse aux traitements médicamenteux chez les patients recrutés dans les essais cliniques. Dans leur analyse, le facteur *interest-activity* (représentant les symptômes de perte d'intérêt, asthénie/fatigabilité, aboulie, inactivité) apparaît comme un prédicteur important de l'échec de traitement dans les deux cohortes.

Suivant cette approche, Chekroud et al. se servent aussi de données d'essais cliniques (questionnaires d'auto-évaluation pré-thérapeutiques), incluant 1949 sujets dans la cohorte d'entraînement (STAR*D, comme Uher et al.), et presque 500 pour la validation (COMED). Mais au lieu d'avoir une hypothèse a priori sur la structure des facteurs prédictifs de l'efficacité du citalopram (à la manière d'Uher et al. avec leurs neuf dimensions de symptômes), ils extraient directement les schémas prédictifs pertinents des données cliniques individuelles des patients, sélectionnant ainsi 25 variables d'intérêt. Les figures 20 et 21 résument la méthodologie d'étude et les résultats principaux, et en soi illustrent bien comment la communication et la visualisation des informations peut améliorer leur intelligibilité. Pour autant, la performance du prédicteur est plutôt faible, même si elle semble stable. La plupart des mesures de performance (précision, sensibilité, spécificité, valeurs prédictives négatives et positives) oscillent autour de 60 %, sans dépasser 65 %, tant dans la cohorte d'entraînement qu'en validation. Les auteurs constatent justement que leur modèle n'est pas assez puissant pour justifier un rejet du citalopram chez un patient prédit comme non-répondeur. Bien qu'elle puisse soutenir la robustesse des résultats, la rigueur méthodologique ne mène pas forcément aux résultats positifs espérés. Mais il est tout aussi important d'avoir un socle de données négatives fiables, pour orienter les futures recherches vers des questions judicieuses, au lieu de poursuites futiles.

Abordant la question différemment dans une seconde étude, Chekroud et al. explorent l'effet des antidépresseurs sur des groupes de symptômes spécifiques, plutôt que d'évaluer une mesure de rémission globale (Adam M. Chekroud et al., 2017). Ils définissent des clusters de symptômes empiriquement, grâce à un regroupement hiérarchique des différents items d'échelles de dépression recueillies avant traitement, chez des patients (essai clinique STAR*D, puis reproduit dans COMED et testé dans neuf autres essais). Ils observent d'une part que différents antidépresseurs n'ont pas la même efficacité sur un cluster de symptômes donné, et d'autre part que les différents clusters ne répondent pas de la même manière aux antidépresseurs. En revanche, leur tentative d'intégration de ces clusters dans des modèles prédictif d'efficacité s'avère plus difficile, avec une grande variabilité des modèles ou un effet clinique modeste. L'analyse non-supervisée (regroupement hiérarchique), avec sa part d'arbitraire, peut être plus difficile à maîtriser, et on peut voir ici comment les mêmes données (STAR*D et COMED), analysées et représentées différemment, peuvent montrer des résultats variables d'une étude à l'autre.

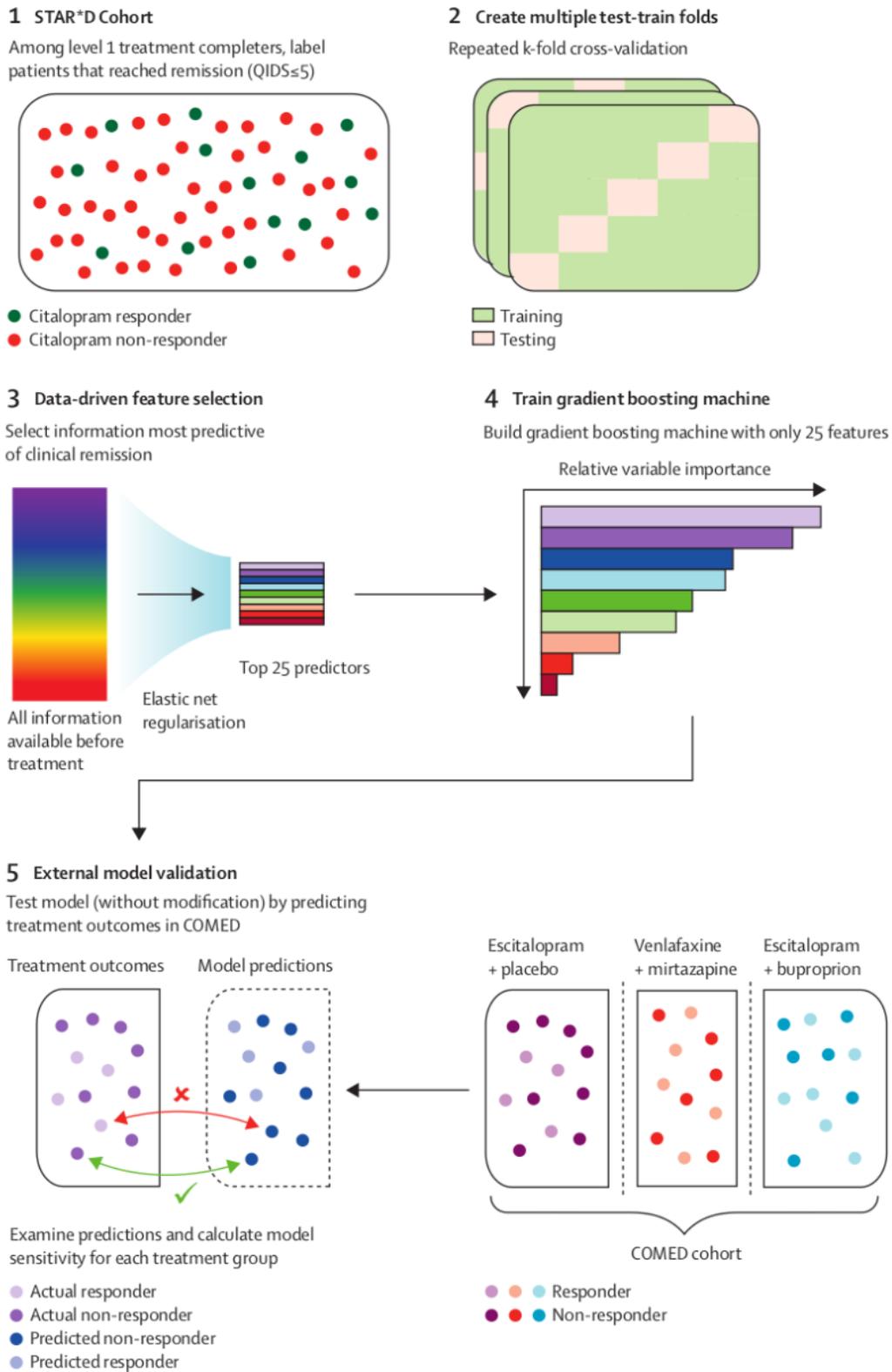


Figure 20 : Protocole de traitement des données et d'entraînement du modèle de prédiction (Chekroud et al., 2016)

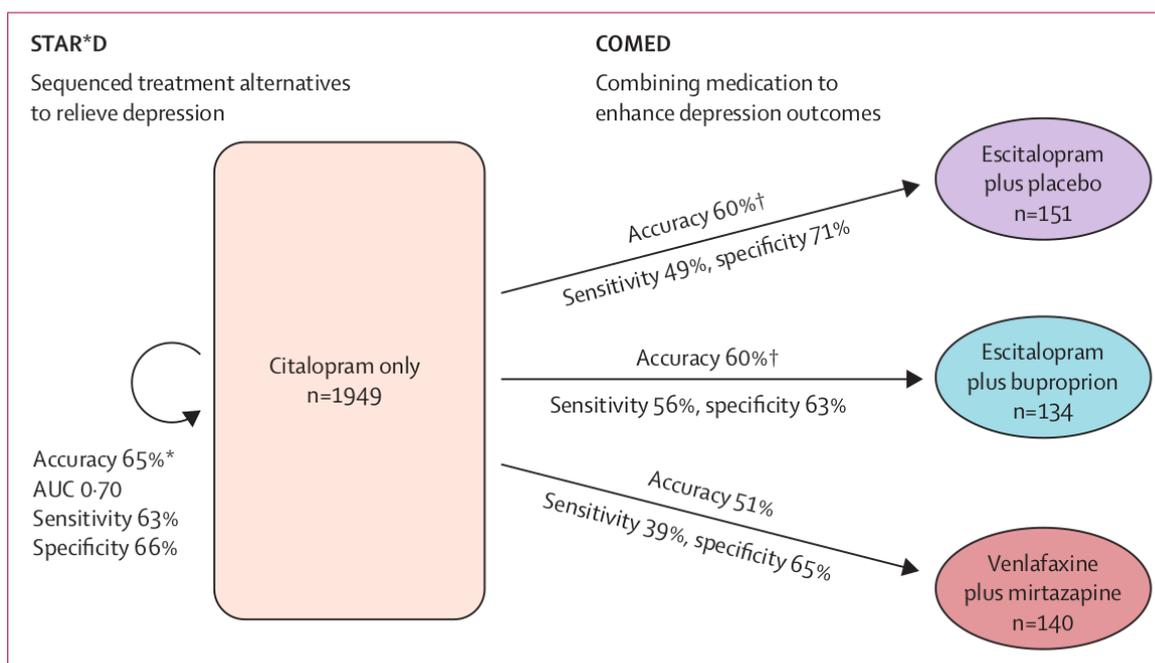


Figure 21 : Résultats principaux de performance sur les données d'entraînement et cohortes test
(Chekroud et al., 2016)

Pour revenir à une autre étude exemplaire, Koutsouleris et al. analysent les données du European First Episode Schizophrenia Trial (EUFEST), essai clinique multicentrique randomisé avec un suivi longitudinal d'un an examinant l'efficacité de cinq traitements antipsychotiques chez des patients présentant un premier épisode psychotique (Koutsouleris et al., 2016). Ils cherchent un modèle prédictif d'échec de traitement, à partir de données socio-démographiques, cliniques et neurocognitives obtenues chez 334 patients. Leur schéma d'analyse, avec deux niveaux imbriqués de validation interne (validation croisée, et *leave-site-out validation*), ainsi qu'une validation externe sur 108 patients, aboutit à la sélection de variables et de deux modèles (SVM) donnant un pronostic à quatre semaines et à 52 semaines.

La valeur prédictive positive du modèle à quatre semaines sur la cohorte de validation externe est de 92 %, montrant une excellente performance pour prédire une non-rémission après quatre semaines de traitement. Cependant, le taux de rémission à cette date est de 33 %, alors qu'il atteint 77 % à 52 semaines. Le modèle prédictif à 52 semaines (*leave-site-out validation*) montre, en effet, une inversion des performances pour le pronostic d'échec et de succès (respectivement, valeur prédictive positive basse, et négative plus élevée), alors que les mesures de sensibilité, spécificité, précision et nombre à diagnostiquer restent semblables. Pouvoir prendre en compte la temporalité du processus de rémission est donc d'une très grande importance. Les variables les plus

pertinentes pour la prédiction d'un échec de traitement sont le statut de chômeur, le niveau d'activité en journée, les besoins relationnels insatisfaits, la détresse psychique, et un faible niveau de scolarité, soulignant, semble-t-il, la place de l'insertion sociale dans la rémission d'un premier épisode psychotique. La prédiction de non-rémission refléterait aussi d'autres mesures de morbidité, comme la persistance des symptômes, une faible qualité de vie, la ré-hospitalisation et le manque d'adhérence au traitement. Le souci, au cœur de cette étude, de la pertinence clinique et d'un outil généralisable et applicable, font l'intérêt de ses résultats prometteurs. On regrette, cependant, l'absence d'indication sur un partage ouvert des modèles prédictifs.

4. Suivre

Le manque de transparence est plus visible dans le domaine du suivi ou du traitement, où la confrontation entre intérêts scientifiques, médicaux et commerciaux est plus évidente. Le paysage est en effet dominé par des logiciels commerciaux, comme en témoigne une comparaison entre la littérature médicale et un parcours des différents vendeurs d'applications mobiles, recensant une trentaine de publications scientifiques concernant les logiciels, mais plus d'un millier d'applications commerciales pour la dépression (Martínez-Pérez, de la Torre-Díez, & López-Coronado, 2013). De plus, il existe un grand nombre d'applications visant à améliorer le « bien-être » psychique sans désigner explicitement une pathologie psychiatrique (Anthes, 2016). Ce type de logiciel n'est pas soumis à la réglementation concernant les dispositifs médicaux, même si en pratique des patients pourraient essayer de s'en servir à visée thérapeutique et donc s'exposer à des outils à l'efficacité non-prouvée et dont la nocivité n'est pas exclue. La Haute Autorité de Santé (HAS) a un référentiel de bonnes pratiques sur les applications et les objets connectés en santé (Haute Autorité de santé, 2016), mais en réalité, il est difficile d'assurer la qualité et la conformité des produits téléchargeables, ou d'améliorer leur développement (Bhugra et al., 2017; Fiveash, 2017; Nicholas, Fogarty, Boydell, & Christensen, 2017). Plusieurs revues soulignent le manque d'études d'efficacité concernant les logiciels de santé mentale, et la difficulté à tirer des conclusions à partir d'études préliminaires, le manque de données sur leur coût et leur rentabilité et sur la généralisation possible de leur adoption (Bakker, Kazantzis, Rickwood, & Rickard, 2016; Grist, Porter, & Stallard, 2017; Hollis et al., 2017). Deux méta-analyses sur des

applications visant les symptômes anxieux et dépressifs concluent à des effets modérés (dépression) à faibles des applications par rapport à un comparateur inactif, mais d'effet incertain par rapport à un comparateur actif, en sachant que les populations incluses comprenaient aussi des sujets infra-cliniques, et pas seulement des pathologies avérées (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017).

Les premiers essais de traitements ou de suivis numériques étaient basés sur des sites internet, avec une efficacité et une adhérence variables selon les modalités, mais certains résultats prometteurs, notamment en TCC avec intervention humaine (Donker et al., 2015; Seyffert et al., 2016). Toute approche numérique n'est cependant pas synonyme de Big Data, ces premières itérations cherchant plutôt à améliorer l'accessibilité et la flexibilité du suivi, en mettant à disposition en ligne un protocole de suivi bien formalisé. Les applications mobiles sont arrivées par la suite, exploitant la mobilité accrue du téléphone portable, ses possibilités d'enregistrement passif de données et une meilleure interface utilisateur. Même si elle ne concerne pas toutes les applications mobiles, la dimension Big Data devient alors plus claire : la possibilité d'un *monitoring* continu, et de *feedback* intelligent en temps réel, grâce à la collecte et l'analyse de données individuelles massives. Malgré cette ambition, la plupart des logiciels non-commerciaux utilisant une approche Big Data qui font l'objet d'une publication scientifique/médicale, sont encore au stade d'études de faisabilité, avec peu d'essais d'efficacité (commentaire dans Bhugra et al., 2017; Roess, 2017, et pour citer quelques études, Ben-Zeev et al., 2014; Biagianti, Schlosser, Nahum, Woolley, & Vinogradov, 2016; Bucci et al., 2018; Schlosser et al., 2018; Torous et al., 2014).

De même, il existe des collaborations entre des éditeurs d'applications commerciales, ou des start-ups, et des institutions publiques ou des universités, pour tester certains logiciels (« Home - NHS Apps Library », s. d.; « Product Platform Development & Pipeline - Mindstrong Health », s. d.; « Science – Medibio », s. d.; « Woebot - Your charming robot friend who is ready to listen, 24/7 », s. d.). La plupart des études sont toujours en cours, les résultats déjà publiés encore très préliminaires avec des méthodologies critiquables (Dagum, 2018; Fitzpatrick, Darcy, & Vierhile, 2017), ce qui n'empêche pas les start-ups d'utiliser ces études et ces collaborations comme outil promotionnel en affichant une confiance excessive en ces résultats. Le besoin de préserver la confidentialité de technologies brevetées va à l'encontre d'un partage ouvert des algorithmes et des données utilisées pour leur développement, entravant ainsi un

examen éclairé de leur fonctionnement. L'article publié par Paul Dagum, fondateur de la start-up Mindstrong Health avec Thomas Insel, ancien directeur du NIMH, illustre cette tension entre intérêt scientifique et commercial (Dagum, 2018). Mindstrong tente d'identifier un « phénotype numérique » (*digital phenotyping*), par des techniques d'apprentissage automatique ou d'intelligence artificielle, c'est-à-dire des marqueurs numériques passifs (les frappes, appuis et glissements de doigts sur l'écran d'un smartphone) reflétant fidèlement l'état cognitif de l'utilisateur, le but étant d'avoir un suivi continu capable d'alerter l'utilisateur au glissement vers un état pathologique avant même qu'apparaissent des signes cliniques. Introduisant le concept derrière Mindstrong, et des résultats préliminaires de son utilisation analysés par la start-up elle-même, cette étude présente les résultats de manière peu informative, même si l'auteur indique que l'algorithme et les données sont disponibles sur demande raisonnable (« on reasonable request »). Medibio cherche aussi une modalité de suivi continu avec des marqueurs passifs pour détecter la survenue de maladies psychiatriques et améliorer la gestion de symptômes. L'algorithme repose sur des mesures physiologiques (fluctuations circadiennes de la fréquence cardiaque), nécessitant le port d'un appareil supplémentaire (« Science – Medibio », s. d.). Le site de la start-up mentionne plusieurs études cliniques en cours non publiées, mais fait déjà état de la supériorité de son outil par rapport à certaines échelles utilisées en recherche.

D'autres start-ups orientent leur activité vers l'offre de suivi thérapeutique en ligne, par messagerie ou vidéo-conférence. Certaines, mettant en place des plateformes de rencontre entre utilisateurs et thérapeutes, analysent les données des utilisateurs par des techniques d'apprentissage automatique pour, selon elles, mieux répondre à leurs besoins et mieux les apparier avec un thérapeute adéquat (« Ginger.io », s. d.; Talkspace, s. d.). Par ailleurs, les outils d'intelligence artificielle qu'utilise Talkspace sont nuls autres que ceux commercialisés par Watson Health d'IBM (Buhr, 2015; IBM News Room, 2015). En l'absence de publications scientifiques ou d'informations plus détaillées, il est impossible de juger le fonctionnement de ces deux start-ups. À la différence des deux plateformes précédentes, Woebot¹⁰ met en lien des utilisateurs humains avec un chatbot, un agent conversationnel autonome, qui selon ses créateurs, aurait été entraîné à appliquer des

10 Le nom est un jeu de mots, presque un calembours, plutôt astucieux, entre « robot » et « woe » signifiant malheur ou infortune en Anglais. Le remplacement du « r » par un « w » dans la prononciation rappelle un défaut d'énonciation infantile commun en Anglais, et donne à ce nom un caractère amusant et attendrissant, qui correspond parfaitement à la « personnalité » créée pour ce chatbot.

principes de TCC pour entretenir des conversations en apportant un soutien émotionnel (Fitzpatrick et al., 2017; « Woebot - Your charming robot friend who is ready to listen, 24/7 », s. d.).

Au-delà de l'opacité des algorithmes, ces initiatives soulèvent des questions autour de la maîtrise et de la réutilisation de données. Les données d'enregistrement et les échanges ayant lieu sur ces plateformes peuvent être sauvegardés par les fournisseurs du service, et utilisés pour optimiser les algorithmes, ou partagés avec des entités tierces. La lecture des conditions d'utilisation et de la politique de confidentialité d'un service comme celui de 7 Cups of Tea, application mobile de soutien et de thérapie par messagerie, donne une idée de l'étendue possible d'accès et d'utilisation des données, sans réellement donner d'éclairage précis sur les partenaires tiers, les modalités de transfert et d'utilisation, la monétisation ou non de ces données (Dobbs, 2017; « Online Therapy & Free Counseling, Someone To Talk To | 7 Cups », s. d.) etc. Sachant que la majorité des utilisateurs, qui plus est une personne en détresse, ne lira pas ces documents, devrait-il y avoir des restrictions et des protections supplémentaires dans le cadre de plateformes qui affichent un service de soin ou d'accompagnement ? Avec des données aussi sensibles, une réutilisation qui ne serait pas dans l'intérêt direct de l'utilisateur, ou dans l'intérêt public, est-elle légitime ou est-ce une forme d'exploitation, même quand ces données sont anonymisées ? Comment et qui définira cet intérêt ? Une plateforme pourrait, par exemple, considérer que l'intérêt de l'utilisateur est synonyme de sa satisfaction, et choisir de mesurer cette satisfaction par le taux de retour à la plateforme. Un algorithme optimisant ce taux de retour ferait tout pour maintenir les utilisateurs sur la plateforme ou les faire revenir... même peut-être en suggérant un suivi sous-optimal pour que l'utilisateur n'aille pas mieux trop vite ? Ce scénario est caricatural, mais sert à illustrer les risques d'un manque de transparence des méthodes et des algorithmes, de la confusion des objectifs et des intérêts, lorsque le sujet en jeu est la vulnérabilité de personnes humaines, que ce soit dans le soin médical, ou l'accompagnement de la détresse. Nul besoin d'une intention malveillante pour faire du mal, un choix peu judicieux de données d'entraînement, d'algorithmes ou de critères d'optimisation suffirait pour causer du tort en toute bonne foi et en toute ignorance¹¹.

11 Cet argument est développé par la mathématicienne et *data scientist* Cathy O'Neil dans son livre récent, *Weapons of Math Destruction* (O'Neil, 2016a), qui documente les manières dont le Big Data peut renforcer les inégalités sociales. Une adaptation du livre, examinant l'utilisation des algorithmes prédictifs et du Big Data dans le monde du travail et des ressources humaines, est publié sous la forme d'un essai

Le questionnement autour des données apparaît sous un autre angle dans le cas d'une start-up comme Spring Care. Fondée par certains des auteurs de l'étude de Chekroud et al. (Adam Mourad Chekroud et al., 2016), elle semble offrir un bouquet de services en santé mentale, ciblant des entreprises et leurs employés comme clients (« Employers - Spring Health », s. d.). Les prises en charge proposées ne sont pas décrites en détail, mais il est fait mention d'apprentissage automatique, d'algorithmes propriétaires et d'études publiées, qui suggèrent que les modèles développés dans ces études seraient à l'oeuvre ici (Adam M. Chekroud et al., 2017; Koutsouleris et al., 2016). Dans quelle mesure les données personnelles de santé, même anonymisées, qui sont le fruit de fonds publics – les essais STAR*D et COMED ont tous deux été financés par le NIMH – devraient-elles pouvoir servir à des initiatives d'ordre commercial ? Toutes les réutilisations, suffisamment respectueuses de la vie privée, peuvent-elles être autorisées ? Ou devrait-il y avoir un cadre particulier pour les réutilisations donnant lieu à un dispositif propriétaire et/ou commercialisé : gratuité de la licence du produit développé pour les organismes publics, par exemple ? Il y a quelque chose qui dérange dans ce type d'utilisation, peut-être en raison de la nature essentiellement intime des données de santé, et de surcroît, celles concernant le psychique, de la fragilité qu'elles exposent. Mais aussi peut-être parce qu'elles sont le fruit d'une perte d'autonomie, tant au sens fonctionnel qu'au sens moral. Le malade se trouve dans une situation de besoin, de nécessité, quand il demande des soins et doit se dévoiler physiquement et personnellement. Ces données apparaissent dans une situation de limite à sa liberté évidente¹². Ce déséquilibre ne devrait-il pas inviter à une certaine réserve, une certaine pudeur, une utilisation « consciencieuse », pour ne pas porter atteinte à la dignité de la personne, ne pas diminuer son humanité ? La désincarnation des données peut faire oublier la nécessité de considérer la personne toujours comme une fin, et non seulement un moyen (au sens de l'impératif kantien). On peut voir la non-commercialité du corps humain (Comité consultatif

dans *The Guardian* (O'Neil, 2016b), et donne un exemple très pertinent ici, de la manière dont une personne avec un trouble psychiatrique peut faire l'objet d'un profilage discriminatoire, dans le processus de demande d'emploi et de recrutement, sans avoir à révéler son diagnostic, mais par le simple recoupement de données.

12 L'argument pourrait être fait que compte tenu de l'hégémonie des GAFAM sur les technologies intimement liées au fonctionnement de nos sociétés modernes, cette limite à la liberté et cette position tributaire ne concerne pas que les malades. Certes, mais on peut reconnaître qu'il existe des degrés différents de perte d'autonomie, qui invitent à une vigilance plus accrue. La situation des détenus, et leur condition d'entrave à la liberté qui nécessite un regard juridique spécifique, en est un autre exemple.

national d'éthique, 2018), avec la gratuité du don d'organes, comme découlant en partie de ce principe, et on conçoit bien le besoin de maintenir un don libre et gratuit pour les organes. Ne pourrait-on pas imaginer la non-commercialité des données individuelles de santé ? Ces questions doivent placer une interrogation éthique au cœur de la politique et de l'économie, en tant que ces domaines sont aux prises avec la montée du numérique, et inversement, ces questions replacent le politique et l'économique au cœur de la médecine, en éveillant une réflexion sur la place du pouvoir, l'allocation de ressources et la citoyenneté dans la pratique médicale.

Malgré le manque de données sur leur rentabilité, l'allocation de ressources est un des arguments centraux pour le développement de suivis sur supports numériques, automatisant certains aspects pour réduire le besoin d'intervention professionnelle humaine, et rendant d'autres aspects plus disponibles géographiquement et temporellement (Bhugra et al., 2017), afin de faire face à la distribution inégale de l'offre de soin, et le changement de la nature des besoins. Ainsi, la plupart des applications et plateformes présentées plus haut répondent à une demande correspondant davantage à des troubles anxio-dépressifs modérés et des troubles de l'adaptation, ou même à une recherche de « bien-être » (*well-being*) psychique. La conception d'outils pour le suivi de maladies psychiatriques sévères semble plutôt avoir lieu au sein de départements universitaires et hospitaliers spécialisés. À titre illustratif, l'on peut présenter ici trois applications qui font l'objet de publications d'essais préliminaires, PRIME (personalized real-time intervention for motivational enhancement) et CrossCheck pour la schizophrénie et les troubles psychotiques, et A-CHESS (Addiction-Comprehensive Health Enhancement Support System) pour la dépendance à l'alcool.

CrossCheck, présenté précédemment (voir sous-chapitre II.1. p. 60, Ben-Zeev et al., 2017), est en cours d'évaluation dans un essai clinique randomisé, mais cinq cas cliniques de patients ayant subi une hospitalisation, sont publiés dans cette étude. Par une surveillance passive continue, en milieu « naturel », de marqueurs comportementaux et des questionnaires d'auto-évaluation (*ecological momentary assessments*) à intervalles réguliers, ne nécessitant pas de consultation avec un soignant, les auteurs espèrent pouvoir détecter les signes avant-coureurs d'une décompensation, afin de prévenir une hospitalisation. Les observations tirées des cinq patients suivis pendant un an montrent une grande variabilité de signes précédant une éventuelle hospitalisation, certains patients les identifiant clairement dans les questionnaires d'auto-évaluation, d'autres manifestant plutôt des modifications de comportement ou d'usage du téléphone. De même, ces

modifications comportementales vont dans le sens d'une augmentation ou d'une diminution d'activité selon le patient. Il est clair qu'il n'existe pas un schéma-type chez tous les patients de signes avant-coureurs d'une décompensation, et que la tendance serait plutôt à des modifications individuelles qui pourraient être détectées par un suivi longitudinal prolongé. La durée du suivi, le nombre d'hospitalisations et de patients, ne permettent cependant pas de conclure à des schémas individuels stéréotypés de décompensation identifiables chez chacun des patients. Les auteurs reconnaissent également que la validité des données enregistrées sont très dépendantes des habitudes d'usage du patient, car une personne qui n'aurait pas en permanence son téléphone à proximité aurait des données moins représentatives de son activité quotidienne.

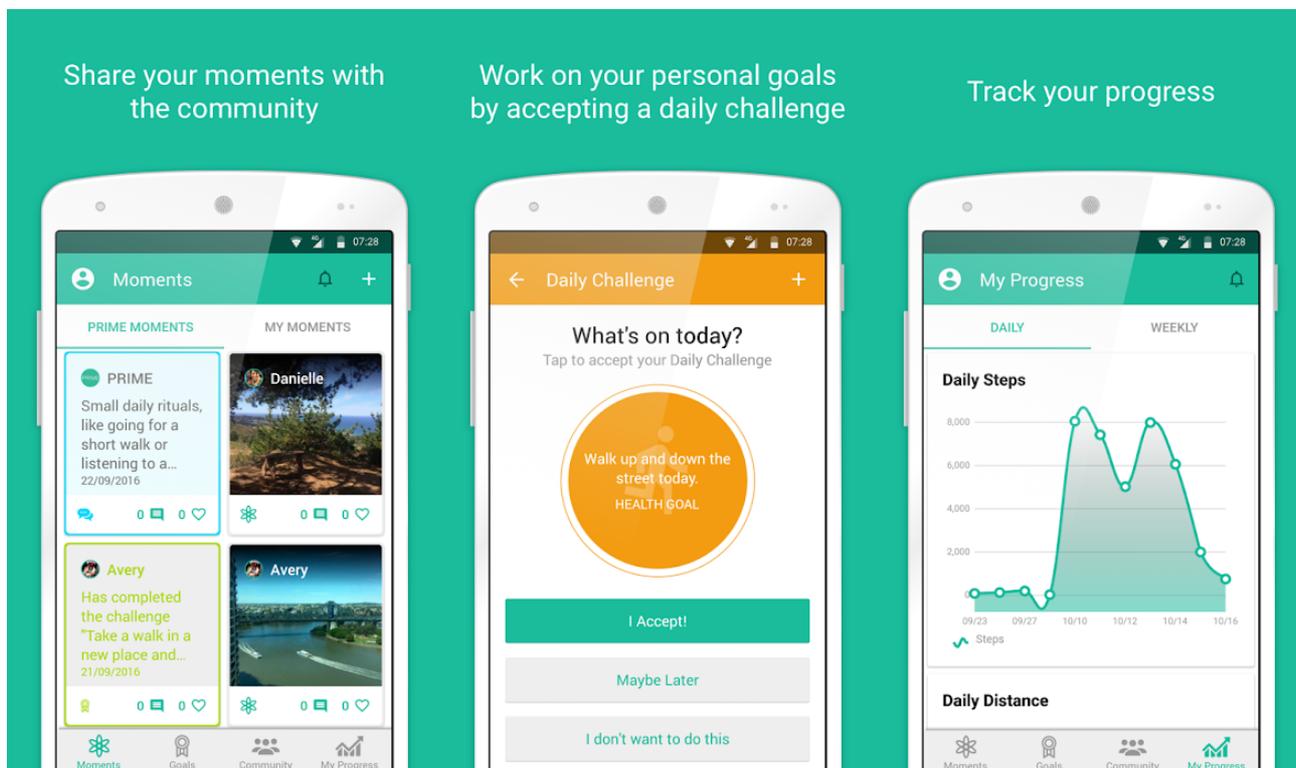


Figure 22 : Captures d'écran de l'application PRIME
(Schlosser et al., 2018; « UCSF PRIME - Apps on Google Play », s. d.)

PRIME (Figure 22) a une vocation plus clairement thérapeutique, proposant une sorte de remédiation cognitive à travers un suivi motivationnel basé sur des principes de TCC, la définition de buts personnels à atteindre avec une décomposition en plusieurs étapes et un suivi personnalisé, et l'accès à une plateforme de messagerie entre pairs, et avec des soignants (Schlosser et al., 2018). L'essai randomisé inclut 43 patients en début de maladie (diagnostiqués depuis moins de cinq ans), 22 utilisant PRIME pendant douze semaines, 21 sur liste d'attente. Il a été réalisé entièrement à distance, les évaluations pré

et post-interventionnelles se faisant par télé-consultation, les patients recrutés provenant de différentes localisations aux États-Unis, au Canada et en Australie. L'évaluation à douze semaines, puis trois mois après l'arrêt de l'intervention, montre quelques améliorations chez les utilisateurs de PRIME sur des mesures de motivation et de comportement pro-social, mais pas de modifications de la symptomatologie, de la qualité de vie ni du fonctionnement global. L'application semble bien acceptée par les utilisateurs, mais doit faire ses preuves dans un essai plus long, avec un plus grand effectif et un critère de jugement principal plus cliniquement significatif. Contrairement à CrossCheck où une approche Big Data est évidente, elle l'est moins pour PRIME, si ce n'est la démonstration d'un essai clinique entièrement réalisé à distance, l'enregistrement de données d'utilisation, et l'accès à un réseau social et des mesures personnalisables et personnalisées.

Une autre application, Actissist, pour les patients psychotiques en début de maladie, fait l'objet d'une étude préliminaire, avec un comparateur actif, l'application ClinTouch, qui est une application de surveillance des symptômes par l'administration de questionnaires d'auto-évaluation sur smartphone, à intervalles réguliers (Bucci et al., 2018). Actissist en revanche, supplémente l'auto-évaluation par des conseils basés sur des principes de TCC, proposés par l'application en fonction des problèmes identifiés par le questionnaire. Elle donne aussi accès à des ressources d'information sur différents aspects de la maladie ou sur les difficultés que peuvent rencontrer les patients, ainsi que des instructions pour des exercices de relaxation et de méditation, et des vidéos motivationnelles. L'application a été conçue en collaboration entre des malades, des soignants et des institutions publiques. L'étude randomisée en simple aveugle inclut 36 patients (24 sur Actissist, 12 sur ClinTouch) et trouve que l'application est bien acceptée et utilisée par les patients (mais les patients avaient des incitations financières pour l'utilisation). Bien que ce ne soit pas le critère de jugement principal, les auteurs retrouvent également une supériorité d'Actissist par rapport à ClinTouch sur des mesures d'amélioration clinique et de fonctionnement général.

L'application A-CHESS fait l'objet d'une étude plus conséquente que les précédentes, avec 349 patients recrutés dans des établissements de soins de réadaptation alcoolique (Gustafson et al., 2014). Randomisés à leur sortie (179 recevant le traitement habituel et 170 recevant un smartphone avec l'application A-CHESS en plus du traitement habituel), ils ont été suivis pendant huit mois d'intervention, puis quatre mois supplémentaires après l'arrêt de l'intervention. L'application offre un accès à des contenus

fixes, comme des aides à la relaxation, des informations sur l'addiction ; d'autres contenus plus dynamiques comme des alertes pour des réunions d'Alcooliques Anonymes, ou pour d'autres évènements d'intérêt dans la communauté ; et des fonctionnalités interactives comme un bouton panique, un message alertant le patient si les coordonnées GPS situent le patient près d'une zone à risque (un bar par exemple), des questionnaires d'auto-évaluation hebdomadaires, et des services de messagerie et de forums entre pairs et soignants. Le critère de jugement principal, le nombre de jours de consommation à risque évalué à quatre, huit et douze mois, se trouve significativement amélioré chez les utilisateurs de A-CHESS (1,39 jours contre 2,75 jours), avec un lien possible entre l'intensité d'utilisation de l'application et l'effet protecteur. En revanche, il n'y a pas de différence entre les deux groupes sur des mesures secondaires de conséquences adverses d'alcoolisations (perte d'emploi, difficultés financières, accident etc).

Ces applications semblent pour le moins être bien acceptées par leurs utilisateurs, même si les cohortes sont petites et d'autres études ne retrouvent pas la même satisfaction (Mueller et al., 2018; Nicholas et al., 2017). À la différence des applications commerciales, leurs procédures sont publiées ouvertement, et ne comportent pas pour l'instant de composante en apprentissage automatique. Le maintien d'un lien humain semble être une caractéristique importante de ces applications, avec la possibilité de contacter un soignant par messagerie, et les fonctionnalités de réseaux sociaux et de forums entre patients. La supériorité de ces interventions par rapport au « traitement habituel » (regrettablement non défini) reste encore à établir, les études étant très préliminaires et ne montrant pas d'effets sur des mesures d'importance clinique et fonctionnelle pour les patients. Dans des conditions où l'offre de soins est limitée (distance géographique, manque de soignants ou de thérapies spécialisées), ces propositions peuvent être des solutions intéressantes, si elles démontrent leur efficacité. Mais la comparaison devrait-elle se faire avec la situation de pénurie, ou le cas de figure idéal ? Dans notre organisation sectorielle des soins psychiatriques – bien mise à mal depuis plusieurs années (Béguin, 2018a, 2018b; Editorial, 2018) – l'insertion des soins au sein d'un territoire, leur ancrage dans un espace physique et non virtuel (Coupechoux & Delion, 2014), sont des principes fondamentaux, afin de garantir la continuité thérapeutique par une proximité géographique, une distribution équitable de la qualité des soins, mais aussi de s'inscrire dans la cité, d'être un noyau de possibilités citoyennes¹³. L'élimination de cet attachement à un lieu a-t-il un prix ? En termes strictement financiers, il serait important de

13 Tout comme l'école publique du reste, qui doit faire naître cette citoyenneté.

comparer le coût du développement, du maintien et de l'opération de telles plateformes, à celui d'alternatives, comme le financement de postes et de formations d'infirmiers et d'éducateurs suffisants, pour pouvoir déployer des accompagnements comparables¹⁴. Mais peut-on réellement tout comparer : une visite à domicile, un accompagnement physique pour une démarche angoissante, un système automatisé de suivi et d'alerte qui pourrait anticiper une décompensation ? Il est tentant de penser que ces différentes pratiques seraient en complémentarité, en synergie même, l'une soutenant et renforçant l'autre, au bénéfice du patient. Mais l'impératif d'allocation optimal des ressources risque de mettre en péril une co-construction du soin humain et numérique, surtout si l'optimisation est biaisée envers des mesures d'efficacité gestionnaire, plutôt qu'une reconnaissance d'une qualité humaine de vie, certes un peu plus difficile à mesurer.

Élargissant le regard de l'individuel vers le collectif, on peut voir des utilisations de Big Data pour d'autres types de suivi, comme la pharmacovigilance, ou l'étude des tendances de prescription et de consommation de médicaments. Ces classiques de la santé publique peuvent être alimentés par de nouvelles sources de données qui pourraient donner des indications sur les utilisations en « vie réelle », ou donner accès à des informations plus difficiles à obtenir dans des bases administratives. Ainsi, une étude de réseaux sociaux sur l'usage du méthylphénidate a pu repérer des discussions sur les effets secondaires du médicament, mais également sur différents types de mésusage (Foulquié et al., 2018). Une autre étude utilise également des données de réseaux sociaux pour détecter les différents usages hors AMM (autorisation de mise sur le marché) des antipsychotiques (Schück et al., 2018). L'analyse des données de remboursements « classiques » par des techniques d'apprentissage statistique peut aussi permettre de tirer des informations plus fines de ces données, comme dans une étude qui identifie quatre trajectoires temporelles différentes de consommation d'anxiolytiques chez les nouveaux consommateurs de plus de 50 ans, suggérant que les difficultés d'arrêt des anxiolytiques diffèrent selon le profil du consommateur (Verger, Mmadi Mrenda, Cortaredona, Tournier, & Verdoux, 2018). Enfin, Qntfy, une entreprise développant des outils de *data science* pour l'analyse de données comportementales et psychologiques, publie de nombreuses études sur l'analyse de réseaux sociaux pour le dépistage de troubles psychiques, notamment une étude récente sur le dépistage d'idées suicidaires par l'analyse des réseaux sociaux en NLP (Coppersmith, Leary, Crutchley, & Fine, 2018). Travaillant sur

14 La rentabilité des interventions en ligne fait l'objet d'études, mais qui sont assez hétérogènes et où le comparateur est généralement la liste d'attente (Donker et al., 2015).

plusieurs collaborations, dans le secteur privé (avec 7 Cups of Tea, par exemple), universitaire et associatif, Qntfy semblerait avoir une approche favorisant un esprit de recherche, et affiche une volonté d'aborder les interrogations éthiques que soulève son travail (Coppersmith et al., 2018; Qntfy Inc, s. d.). Serait-ce une démarche sincère et intentionnelle de conciliation entre poursuites commerciales et intérêts publics, de mise au travail de ces différentes facettes qui composent le Big Data en psychiatrie ?

III – Outil révolutionnaire ou dangereux mirage ?

« (...) If he would leave the self alone,
Apollo's welcome to the throne,
Fasces and falcons;
He loves to rule, has always done it;
The earth would soon, did Hermes run it,
Be like the Balkans.

But jealous of our god of dreams,
His common-sense in secret schemes
To rule the heart;

(...) And when he occupies a college,
Truth is replaced by Useful Knowledge;
He pays particular
Attention to Commercial Thought,
Public Relations, Hygiene, Sport,
In his curricula.

(...)The goal a populous Nirvana:
His shield bears this device: Mens sana
Qui mal y pense.

(...)In our morale must lie our strength:
So, that we may behold at length
Routed Apollo's
Battalions melt away like fog,
Keep well the Hermetic Decalogue,
Which runs as follows:--

Thou shalt not do as the dean pleases,
Thou shalt not write thy doctor's thesis
On education,
Thou shalt not worship projects nor
Shalt thou or thine bow down before
Administration.

Thou shalt not answer questionnaires
Or quizzes upon World-Affairs,
Nor with compliance
Take any test. Thou shalt not sit
With statisticians nor commit
A social science.

Thou shalt not be on friendly terms
With guys in advertising firms,

(...) Thou shalt not live within thy means
Nor on plain water and raw greens.
If thou must choose
Between the chances, choose the odd; »

« (...) S'il laissait le soi seul,
Apollon serait le bienvenu sur le trône,
Faisceaux et faucons ;
Il adore régner, l'a toujours fait ;
La terre sitôt serait, si Hermès la dirigeait,
Comme les Balkans.

Mais jaloux de notre dieu des songes,
Son bon sens en secret complot
De régner sur le cœur ;

(...) Et quand il occupe une fac,
La Vérité est remplacée par le Savoir Utile ;
Il donne une particulière
Attention à la Pensée Commerciale,
Aux Relations Publiques, l'Hygiène, le Sport,
Dans ses programmes.

(...)Le but un Nirvana très peuplé :
Son armure porte cette devise : Mens sana
Qui mal y pense.

(...) En notre moral doit être notre force :
Donc, que nous voyions à terme
D'Apollon, dérouterés,
Les bataillons fondre comme le brouillard,
Garde bien le Décalogue Hermétique,
Qui va ainsi : --

Tu ne feras point comme plaira au doyen,
Tu n'écriras point ta thèse de docteur
Sur l'éducation,
Tu ne vénéreras point les projets ni
Toi ni les tiens ne vous prosternerez devant
L'administration.

Tu ne répondras point à des questionnaires
Ou des quiz sur les Affaires-Mondiales,
Ni avec compliance
Passeras un test. Tu ne t'assoiras point
Avec des statisticiens ni commettras
Une science sociale.

Tu n'entretiendras point de relations amicales
Avec les gars des firmes publicitaires,

(...) Tu ne vivras point dans tes moyens
Ni d'eau plate et de verdure crue.
Si tu dois choisir
Entre les chances, choisis celle de pendu ;(...) »

Under which lyre : a reactionary tract for the times (Auden, 1946)
Sous quelle lyre : un manifeste réactionnaire pour l'époque (traduit par moi-même)

Une césure poétique en guise d'introduction à ce chapitre peut sembler étrange. Mais ces vers irrévérencieux, écrits par le poète Anglo-Américain W. H. Auden, pour la cérémonie de remise des diplômes de l'université d'Harvard dans l'après-guerre immédiat, sont d'une prescience marquante, par leur portrait si juste des dérives d'une raison réduite à sa technicité, au service de l'ordre et du contrôle exigés par les pouvoirs commerciaux et étatiques. Il convient de rappeler ici l'étymologie commune entre « État » et « statistiques » (dérivant du latin *status*, place, position (Centre National de Ressources Textuelles et Lexicales, s. d.-b), reflétant les racines historiques de cette discipline. La confrontation mythique que met en scène Auden entre deux cultures du savoir n'est pas sans rappeler les tensions qui peuvent exister en médecine, et peut-être même plus en psychiatrie, autour de la définition de ce qui est « valide »¹⁵ : en termes de pratique, de connaissances, de cadre de lecture, de nomenclature, de définitions, entre l'*evidence-based medicine* (EBM) des recommandations, et le maillage de courants et d'approches théoriques en clinique (Grenouilloux, 2017; Kendell & Jablensky, 2003; J.-A. Micoulaud-Franchi et al., 2018; Jean-Arthur Micoulaud-Franchi, Dumas, Quiles, & Vion-Dury, 2017).

L'arrivée du Big Data dans cette tour de Babel ferait-elle figure de révolution, faisant table rase des querelles de cliniciens et introduisant une *pax numerica* ? L'EBM tient, en effet, à la faire sienne, invoquant l'objectivité – l'autorité – de données mesurables et de formules mathématiques, et la possibilité attrayante de « hauts rendements » à « faible coût » (« low-cost, high-return », McIntosh et al., 2016). Mais cette vision réductrice (McIntosh et al., 2016; Murdoch & Detsky, 2013; Oquendo et al., 2012; Passos et al., 2016; Torous & Baker, 2016; Weber, Mandl, & Kohane, 2014), et délibérément simplificatrice quand il s'agit de promouvoir une vision politique par exemple (CAB_Solidarités, 2018; Présidence de la République, 2018a; Thompson, 2018; Villani, 2018), empêche d'appréhender l'incertitude de cette vaste entreprise économique, politique et sociale. En fusionnant le futur de la médecine, et la médecine du futur, avec le Big Data, on néglige d'examiner la confrontation de deux cultures¹⁶ du savoir et de la pratique, pourtant sensiblement différentes, en espérant peut-être substituer la culture de

15 À défaut d'avoir accès à, ou de vouloir définir, ce qui est « vrai ».

16 On comprend ici le terme de « culture » au sens sociologique large de « l'ensemble des traits distinctifs, spirituels et matériels, intellectuels et affectifs, qui caractérisent une société ou un groupe social. Elle englobe, outre les arts et les lettres, les modes de vie, les droits fondamentaux de l'être humain, les systèmes de valeurs, les traditions et les croyances » (« Conférence Mondiale sur les Politiques Culturelles (MONDIACULT): rapport final », 1982).

l'un à la culture de l'autre, pour assimiler l'espoir d'une révolution médicale à la révolution actuelle du numérique.

Les deux chapitres précédents tentaient de poser les prémisses de cet examen. En s'attachant à tracer quelques contours historiques du Big Data, à cataloguer le matériel et les méthodes dont il y est fait usage, à esquisser l'écosystème qu'il habite, on espère faire apparaître, plutôt qu'une définition catégorique, un exposé de la « culture Big Data ». Les différentes instances de rencontre entre cette culture et celle de la psychiatrie scientifique et clinique, décrites au second chapitre, permettent ainsi d'illustrer les décalages et les concordances entre elles par l'abord de la technique, incontournable pour un sujet traitant de technologie. Nous pouvons maintenant essayer de comprendre ce qui advient de cette rencontre entre Big Data et psychiatrie.

1. Cultures et pratiques

L'utilisation du Big Data en psychiatrie, que ce soit à travers des bases de données massives, l'analyse de données haute-dimension multimodales, ou l'utilisation d'apprentissage automatique, montre des pistes de recherche très intéressantes, mais encore très préliminaires, pas assez fréquemment répliquées ou assez bien validées. L'impact actuel sur la pratique clinique est encore moins ressenti, aucun des modèles ou des outils décrits ici n'ayant abouti jusqu'ici à une transformation significative des pratiques. Les auteurs d'une revue très détaillée sur l'apprentissage profond dans les sciences biomédicales arrivent à un constat similaire dans une conclusion lucide qu'il convient de citer :

« Deep learning-based methods now match or surpass the previous state of the art in a diverse array of tasks in patient and disease categorization, fundamental biological study, genomics, and treatment development. Returning to our central question: given this rapid progress, has deep learning transformed the study of human disease? Though the answer is highly dependent on the specific domain and problem being addressed, we conclude that deep learning has not yet realized its transformative potential or induced a strategic inflection point. Despite its dominance over competing machine learning approaches in many of the areas reviewed here and quantitative improvements in predictive performance, deep learning has not yet definitively “solved” these problems. » (Ching et al., 2018)

Ce bilan s'étend au-delà du domaine scientifique, aux différentes applications mobiles et logiciels de santé mentale dont la grande majorité n'ont pas de preuve d'efficacité, ni même de non-nocivité (Anthes, 2016). L'exemple le plus marquant vient non pas de la psychiatrie, mais de l'oncologie, domaine qui affiche pourtant parmi les plus grands succès de l'utilisation du Big Data. Montrant bien que l'identification de cibles cellulaires et de facteurs génomiques ne sont pas un problème équivalent à la formulation d'un diagnostic et d'une décision thérapeutique, le système d'aide à la décision médicale de Watson for Oncology (IBM) a rencontré de nombreuses difficultés lors de sa commercialisation et sa mise en application (propositions erronées ou obsolètes, voire dangereuses, bénéfice faible), malgré le travail colossal qu'a nécessité la conception du système (Ross & Swetlitz, 2017, 2018; Swetlitz & Ross, 2018).

Loin de former une culture monolithique, le secteur numérique est tout de même dominé par une représentation sculptée par le marketing des GAFAs, et l'insertion de logiques commerciales à tous ses niveaux. Cet ersatz d'une culture, auquel les nouveaux-venus du secteur cherchent néanmoins à s'assimiler dans l'espoir d'accéder au succès par contiguïté, véhicule l'image d'un entrepreneuriat effréné alimenté par la perception que l'échec a un faible coût. Il possède son propre credo, celui de la « disruption »¹⁷ (Bower & Christensen, 1995), de l'innovation synonyme de progrès, et du « move fast and break things » (« bouger vite et casser les choses », « Form S-1 - Registration Statement under the Securities Act of 1933 - Facebook, Inc. », 2012). Toute tradition est bonne à briser pour faire place à des idées neuves, toute idée nouvelle bonne à tester sans trop réfléchir : un démantèlement sans remise en question. L'influx important de capital dans ce secteur outre-Atlantique, y compris en bio-technologie et santé numérique (Figure 23, Zweig, Tecco, & Huang, 2018), encourage une promotion exagérée des succès et du potentiel de cette culture technologique.

17 Théorie mise en avant dans les années 90 par C. Christensen, professeur à la Harvard Business School, expliquant pourquoi des entreprises établies font faillite : une petite entreprise disposant de moins de ressources serait en mesure de concurrencer avec succès les entreprises régnantes grâce à une « innovation disruptive », innovation en théorie moins chère, souvent de qualité moindre, mais ciblant une clientèle négligée par les entreprises établies (Christensen, Raynor, & McDonald, 2015). Souvent élevée au rang de dogme dans la culture start-up, les fondements de cette théorie sont critiqués (Lepore, 2014). Autant dire qu'un État qui veut faire de la « disruption », cela n'a pas de sens.



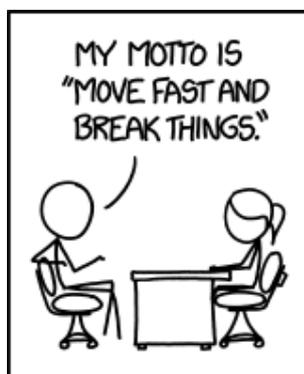
Figure 23 : Investissements récents dans des start-ups en technologies de la santé mentale (Zweig, Tecco, & Huang, 2018)

Autant de tendances qui peuvent trouver résonance avec certaines évolutions regrettables dans le monde de la recherche (notamment, la poursuite acharnée de publications comme fin en soi, qui déforme la science), mais qui se heurtent au final à la réalité du travail clinique, du soin, et du travail scientifique, et à leurs leçons d'humilité quotidiennes. On ose espérer, en effet, que la culture médicale et la culture scientifique tiennent encore à des principes – un sens du devoir et de la responsabilité envers les êtres humains, une recherche de sens au-delà de la simple utilité – qui se situent sur un registre autre que celui de l'efficacité opérationnelle et gestionnaire, de la célébration de « l'exceptionnalisme ». La science et la médecine doivent, certes, s'ouvrir à une interrogation plus vigoureuse de leurs dogmes, de pratiques qui s'ossifient par complaisance, mais en cela, elles sont en droit de demander une approche réfléchie, argumentée, rigoureuse, qui ne perd pas de vue leurs valeurs centrales d'humanité et de rationalité.

Si Watson for Oncology nous montre l'achoppement du secteur technologique face à la clinique, le partenariat entre DeepMind et le Royal Free Hospital Trust du NHS (Royaume-Uni) illustre les méfaits de vouloir appliquer le slogan « move fast and break

things » en santé (entre autres domaines, Figure 24, (« xkcd: Move Fast and Break Things », s. d.). Dans une analyse juridique et éthique, Powles et Hodson interrogent la mise en place de ce partenariat qui a conduit au transfert de données médicales non-anonymisées concernant des millions de patients du Royal Free Trust à des serveurs employés par DeepMind, ostensiblement pour développer un outil d'alerte au risque d'insuffisance rénale aiguë (Powles & Hodson, 2017). Outre la légalité douteuse de ce transfert et la rapidité avec laquelle l'accord s'est établi, les auteurs soulignent les manquements à la déontologie médicale, les patients n'ayant pas été informés du partage de leurs données ni leur accord demandé, et le manque de discussion ou même de regard public sur un accord de cette envergure impliquant un service public et une entreprise privée. Ils questionnent le bénéfice public obtenu, qui semble minime, l'accord ne semblant pas garantir le contrôle public de l'outil développé à partir de ces données.

Figure 24 :
 « Move Fast and Break Things »
 (« xkcd: Move Fast and Break Things », s. d.)



JOBS I'VE BEEN FIRED FROM
 FEDEX DRIVER
 CRANE OPERATOR
 SURGEON
 AIR TRAFFIC CONTROLLER
 PHARMACIST
 MUSEUM CURATOR
 WAITER
 DOG WALKER
 OIL TANKER CAPTAIN
 VIOLINIST
 MARS ROVER DRIVER
 MASSAGE THERAPIST

Ce cas inspire au moins deux réflexions : premièrement, un établissement public ne garantira pas à lui seul la protection de l'intérêt public ni même celui de ses patients, par simple vertu du fait qu'il a une mission abstraite de « service public ». C'est dire qu'une démarche éthique dans ce domaine ne peut être le fait d'une logique réglementaire – logique qui est celle de l'établissement – mais doit se référer à une réflexion collective, donc fondamentalement politique. Ceci nous amène à la deuxième observation, celle de l'asymétrie entre la source des données (hôpital public, patients, individus) et les entités qui détiennent les moyens d'extraire et de traiter ces données, à une échelle qu'on peut

qualifier d'industriel, pour en tirer une valeur ; une asymétrie qui crée des inégalités et fait le terreau d'un système d'exploitation, dans le sens humain et non pas informatique (boyd & Crawford, 2012). Comment empêcher que l'utilisation du Big Data en santé ne renforce ces inégalités, et ne concentre davantage de pouvoir entre les mains des quelques acteurs déjà dominants ? Comment parer à l'érosion de la dignité et de l'autonomie ? On ne pourrait avoir la prétention de résoudre ces questions ici-même, si ce n'est de proposer

que leur discussion ne peut être que l'affaire d'experts ou d'une directive gouvernementale. On commence à voir apparaître, fort heureusement, une délibération publique du numérique en santé, de plus en plus visible dans les Espaces de Réflexion Éthique Régionaux, par exemple, ou lors de la consultation publique en ligne des États Généraux de la Bioéthique 2018 (Comité consultatif national d'éthique, 2018), outils qui amorcent peut-être un processus démocratique, une mise au travail d'un collectif citoyen. Mais le regard décalé de la psychiatrie, que ce soit celui des soignants ou des soignés, manque encore cruellement. Et il ne faut pas oublier que l'exécutif, à grand renfort de déclarations, de commissions et de rapports, bien documentés mais tout de même biaisés (CAB_Solidarités, 2018; Ministère des Solidarités et de la Santé, 2018a, 2018b; Mission Villani sur l'IA, s. d.; Pon & Coury, 2018), n'attend pas forcément l'issue du processus public : il s'empresse d'exécuter, c'est son rôle après tout.

On l'aura donc vu, à la rencontre entre la culture numérique et la culture médicale, il ne faudrait pas que la médecine perde sa route. Parmi les propositions faites pour un « virage numérique » (Pon & Coury, 2018) du système de santé dans le plan « Ma santé 2022 », le terme « éthique » apparaît fréquemment, plutôt comme qualificatif, le rapport s'en tenant à citer quelques principes directeurs, et renvoyant une réflexion éthique, qui n'est certes pas dans ses missions, à d'autres instances et dispositifs. Il insiste sur un engagement fort de l'État pour garantir de la qualité et de la conformité des outils numériques de santé, mais aussi pour le développement d'un socle d'outils de base permettant un accès généralisé et uniforme à ces outils. Cette volonté de favoriser l'initiative publique et de maintenir une veille publique sur le secteur est encourageante, se distinguant de ce qui se fait dans d'autres États. Il reste encore à voir lesquelles de ces propositions seront adoptées et comment elles se traduiront en une politique de santé numérique. De plus, l'implication de l'État ne garantit pas une réelle démocratisation du secteur, pouvant se limiter à l'instauration d'une bureaucratie plus ou moins efficace et égalitaire.

« Homogène », « stabilisé », « généralisé », « standard » (Pon & Coury, 2018), autant de termes qui dénotent un dessein de contrôle sur le développement du numérique en santé. Mais par extension, ne serait-ce pas aussi, avec la promesse d'une meilleure « qualité » et rentabilité du soin, d'une meilleure « sécurité » du soin grâce à sa « traçabilité » (Villani, 2018), une manière de mieux dompter tous ces soignants disparates, surtout ceux exerçant des professions « libérales » ? En maîtrisant les outils de sa pratique ; en fixant peu à peu la forme à travers laquelle il transcrit, communique et

garde en mémoire son travail de soin ; en structurant même jusqu'au langage dont il se sert, à travers une terminologie de référence qui s'imposera facilement grâce aux fonctions d'auto-complétion de texte par exemple, n'est-ce pas la structure de la pensée du soignant, en plus de sa pratique, qu'on essaie de contrôler, au risque d'homogénéiser les façons d'être soignant ?

Curieuse omission pour une démarche qui entend « personnaliser » la médecine avec l'apport de l'apprentissage automatique, de ne pas interroger le « compromis biais-variance » de la relation soignante – si l'on peut dire – la manière dont, à vouloir rendre le modèle trop généralisable, on perd du singulier dans la relation thérapeutique, et on aplatit l'hétérogénéité. Or en psychiatrie, on se demande comment on pourrait soigner sans cette hétérogénéité, que François Tosquelles et Jean Oury décrivent comme étant une nécessité pour construire un collectif de soin autour de patients psychotiques. La question de la « personnalisation » des soins est bien là, dans leurs efforts et ceux de beaucoup d'autres cliniciens, pour concevoir, ou plutôt pour travailler à, une psychiatrie qui ne soit pas déshumanisante : comment penser « des possibilités de faire des systèmes collectifs où l'on puisse vivre d'une façon très personnalisée » (Oury, 2005). Ce travail qui commençait déjà autour de la Seconde Guerre Mondiale, a formé une pratique, un mouvement qui sera nommé mouvement de la psychothérapie institutionnelle (Delion, 2014; Oury, 2007b), et dont peut-être une des conditions essentielles est un processus permanent « d'analyse institutionnelle » (Oury, 2006b), au sens d'une dialectique qui questionne sans cesse le fonctionnement de l'institution et permette de créer les conditions pour que ça « prenne sens » dans la relation au malade, pour qu'il y ait une vraie rencontre, et ainsi une possibilité de soin.

Avant de vouloir faire du neuf, il serait bon de se rappeler que les expériences du mouvement de la psychothérapie institutionnelle ont nourri et guidé les discussions qui donneraient naissance à la psychiatrie de secteur, avec sa vision de la continuité de soins entre l'intra- et l'extra-hospitalier, d'un soin articulé autour du malade et non de l'établissement (Coupechoux & Delion, 2014; Delion, 2014). Sans prise en compte et sans mise au travail d'une logique institutionnelle, le risque est que la médecine personnalisée de demain ne soit qu'un catalogue de procédures sur-spécialisées, à l'efficacité optimisée selon un certain nombre de critères, de préférence biologiques, qui seraient à rechercher ou déjà renseignés par le malade, le tout répondant à une logique consumériste plutôt qu'à un positionnement éthique. Le risque serait de réduire la pratique à la technique.

Il manque justement de la dialectique et de la contextualisation dans ce vaste projet d'intégration du Big Data à la médecine et aux soins, et comme une certaine anosognosie par rapport à ses contradictions internes, ses paradoxes. Certaines assertions, posées un peu trop facilement, doivent être examinées critiquement pour distinguer l'état de fait de la croyance ou de l'espoir, surtout qu'il y va de la justification même de cette entreprise de reconfiguration du système de santé – *des systèmes*, du reste, étant donné qu'elle se poursuit également dans d'autres pays. Comment prendre en compte et gérer, par exemple, la tension entre d'une part une vision qui aspire à améliorer la santé par une systématisation de la détection et de la réduction du risque (la prédiction et la prévention), et d'autre part l'extension systématique du risque de dommages numériques que cette vision implique ? Ces dommages et ces dangers ne sont pas hypothétiques, ils sont multiples et concrets, comme les violations massives de données ayant touché différents secteurs sensibles aux États-Unis, ou le piratage des systèmes informatiques du NHS qui a fortement perturbé son fonctionnement (Marsh, 2017). Ils ont trait à la fois à la concentration importante des données, aux conséquences inattendues d'algorithmes entraînés sur des jeux de données biaisés, à une utilisation peu scrupuleuse des technologies, notamment à des fins de surveillance et de rigidification des inégalités sociales, à la dépendance sur des systèmes complexes mais mal conçus, mal compris ou mal entretenus etc. L'anonymisation et la dé-identification des données – condition souvent suffisante pour ne pas avoir à solliciter le consentement de patients lors d'une réutilisation en recherche, par exemple – ne sont pas un rempart contre les dommages causés par un vol de données, car les possibilités de recoupements entre des jeux de données extensives peuvent permettre de fait une ré-identification des données (Hayden, 2013; Rothstein, 2010). Des travaux existent qui tentent de répertorier les dommages déjà causés, comme le Data Harm Record (« Data Harm Record », 2017) compilé par le Data Justice Lab de l'université de Cardiff (Royaume-Uni).

On essaiera, bien sûr, de se prémunir contre ces dangers, en tirant les leçons des incidents passés, mais on n'anticipera pas toujours la forme qu'ils prendront, et surtout, cela demandera des ressources importantes. À ce propos, et malgré les ressources numériques en santé dont la France dispose déjà (voir sous-chapitre I.1.b. p. 15), les exigences de standardisation, d'interopérabilité, de généralisation, et les intentions d'analyses et de modélisations complexes, demanderont un investissement considérable qui n'aura pas toujours de retour, immédiat ou différé. En effet, on a vu en psychiatrie du moins, combien il était difficile d'obtenir des résultats cliniquement pertinents de ces

analyses. Cela impliquera donc des choix de financement, des redirections de ressources, qui peuvent devenir délétères s'ils ne sont pas soumis à un réexamen permanent des biais de jugement, des partis pris, des dogmes qui les sous-tendent.

Cette discussion ne constitue en aucun cas un rejet des technologies du Big Data ni de leur apport possible à la psychiatrie. Ce sont des technologies très puissantes qui ne cessent, pour l'instant, de se développer et de gagner en fonctionnalités. Il faut néanmoins pouvoir reconnaître, sans rien retirer à la stratégie complexe de ces jeux, que gagner aux échecs ou à Go, ce n'est pas le même processus que de diagnostiquer une schizophrénie, encore moins de soigner un patient schizophrène. D'ailleurs, les applications de l'apprentissage automatique en santé qui semblent les plus prometteuses seraient peut-être celles qui ressemblent le plus aux applications hors du domaine médical où l'intelligence artificielle a rencontré du succès, comme l'imagerie (Ching et al., 2018). Dans d'autres domaines d'applications, en dehors de la pratique clinique, un bénéfice du Big Data pourrait peut-être apparaître plus clairement : en santé publique mais ce n'est rien de nouveau (Chowkwanyun, Bayer, & Galea, 2018) ; en gestion hospitalière et logistique ; et dans l'industrie pharmaceutique, pour la recherche de nouveaux médicaments, de nouvelles indications, ou la meilleure conception des essais cliniques (Hirsch et al., 2017; Sheridan, 2018b).

Ce sont donc des technologies très puissantes, mais leur utilisation judicieuse nécessite une interrogation lucide de leurs limites, à la fois celles qui sont variables et susceptibles d'être surmontées, et celles qui sont irréductibles. Cette interrogation requiert une très bonne maîtrise de leur fonctionnement et de leur application dans le domaine médical/psychiatrique (la technique et sa mise à l'épreuve, la pratique), ainsi qu'une disposition critique, surtout pas cynique, de la pensée. Autant dire qu'il faudrait, pour concevoir une utilisation du numérique en psychiatrie, non pas seulement un mélange des artefacts des deux cultures en question, ou pire encore l'imposition des artefacts de l'une à l'autre, mais surtout une rencontre des personnes habitant ces deux cultures, y compris quelques-unes qui seraient à la frontière peut-être et certaines venant de cultures autres, amenant leurs « appareils à penser » distinctifs avec eux, dans une logique non-hiérarchique, et une volonté de remettre en question, de mettre au travail : quelque chose qui puisse faire du collectif pour reprendre un peu sauvagement Jean Oury (Oury, 2005).

Il serait malavisé de poser ici un verdict, dans un sens comme dans l'autre, et on ne devrait s'en tenir qu'à des constatations alimentant des interrogations et des éléments de réflexion. Ceci invite à une analogie, entre les technologies numériques et celles du

nucléaire – analogie que Cathy O’Neil emploie à si bon effet dans le titre de son ouvrage *Weapons of Math Destruction*, ou *Armes de destruction « math-ives »* (O’Neil, 2016a) – en ce qu’elles touchent toutes deux à une question existentielle, l’une comme possibilité d’annihilation soudaine de l’existence humaine, l’autre comme questionnement de ce qui définit l’humain. Pour l’une comme pour l’autre, leur puissance impose qu’on les déploie avec discernement et humilité.

2. Subversion de la raison

*« La raison du plus fort est toujours la meilleure;
Nous l’allons montrer tout à l’heure. »
Le Loup et l’Agneau, Jean de la Fontaine*

Dire que le Big Data n’a pas encore révolutionné la pratique psychiatrique, ce n’est pas dire qu’il n’a rien changé. Mais c’est une transformation qui semble plus de l’ordre de la subversion, c’est-à-dire une déstabilisation qui n’a pas encore atteint son aboutissement, qui n’a pas encore « fait le tour » contrairement à la révolution, laissant ainsi les choses comme en suspens. Cette subversion concerne le savoir, principalement le savoir scientifique, sa définition, la méthode et les conditions de sa découverte. C’est donc une subversion épistémologique et heuristique qui s’opère par la mise en jeu des données massives et des techniques d’apprentissage automatique. Nous allons maintenant tenter de le démontrer.

Un des postulats en faveur du Big Data en psychiatrie tient à l’objectivité présumée de données mesurables, par opposition à la subjectivité de l’entretien clinique, considérée comme moins fiable. Les données sont perçues comme donnant un accès direct à un objet en soi, sans avoir à passer par les filtres de non pas une mais au moins deux subjectivités et de l’intersubjectivité qui se met en jeu entre elles. Lorsque ces données représentent un phénomène ou un objet physique, leur étude s’affuble de l’autorité des sciences naturelles¹⁸, des sciences expérimentales, même quand il n’y a pas d’expérimentation en jeu. L’observation dont découlent ces données est soutenue par une

18 Au sens de la distinction entre « sciences de la nature » et « sciences de la culture » chez Heinrich Rickert (Ganascia, 2010), les premières étudiant les propriétés objectives du monde, et les secondes les phénomènes sociaux résultant d’activités humaines.

matérialité physique rassurante. Et quand les données correspondent à quelque chose de plus abstrait, les rendre quantifiables permet de retrouver cette objectivité par une référence à une mesure standardisée et chiffrée, donc qui aurait la même valeur pour tout le monde, et ne dépendrait, encore une fois, pas de subjectivités. Ainsi, les données semblent être en accès direct à la réalité, qu'elles objectivent en l'établissant de « fait ». Le terme même de « données » suggère une irréfutabilité : c'est « donné », *it's a given*, cela va de soi. Et poussant l'image plus loin, l'absence de filtre est soulignée par l'expression *raw data*, ou « données crues », pour désigner des données supposément vierges de tout traitement.

On cherche à établir une psychiatrie basée sur des « preuves », personnalisée certes, tout en tendant néanmoins vers cette *evidence-based medicine*, dont la logique correspond largement à celle d'une science naturelle (Ganascia, 2010). Une preuve, c'est quelque chose qui établit un fait, la vérité, la réalité des choses. Or les données ne permettent-elles pas cela justement, par leur objectivité ? Donc par équivalence, on se dit que les données sont ces preuves.

Ceci ne constitue pas encore le savoir. Les preuves, les données, nous donnent des informations, et le raisonnement logique donne un cadre pour établir des relations entre elles (contradiction, non-contradiction par exemple). Mais un savoir, du moins scientifique, est plus qu'une preuve qui concerne une instance particulière. C'est une loi, ou au moins une affirmation que l'on voudrait générale. Comment passer d'une affirmation particulière, découlant d'une instance d'observation particulière, à l'affirmation générale qui constitue le savoir, comment inférer la dernière de la première ? Si on observe cette affirmation particulière se reproduire un grand nombre de fois, ne pourrait-on pas alors la poser en loi générale ? Formulé autrement : est-il justifié d'inférer, à partir d'observations passées répétées, des conclusions sur des observations futures ? Ce type de raisonnement correspond à l'inférence par induction. David Hume, cependant, réfuta l'existence d'une justification *logique* pour le raisonnement inductif, postulant que son usage apparemment répandu est lié davantage à l'habitude qu'à sa validité logique (Popper, 1971, 2002). D'ailleurs, on voit aisément une difficulté pratique : comment sait-on qu'on a observé assez de répétitions pour inférer une loi générale, et qu'il n'y aura pas une observation contradictoire non encore observée qui l'invalidera par la suite ? Autrement dit, une loi générale dépassant (« transcends ») l'expérience, elle ne peut pas être justifiée par l'observation ou l'expérience qui ne peuvent donner qu'un nombre fini de répétitions (Popper, 2002). Karl Popper proposa donc, en réponse au problème

d'induction, que le savoir en sciences naturelles, dont on accepte déjà le postulat qu'il doit être empirique, basé sur l'observation, se dérive de la méthode hypothético-déductive et non de l'induction, par un processus d'essais et d'erreurs, plus précisément, par la réfutation de certaines hypothèses et l'acceptation temporaire d'autres hypothèses pas encore réfutées (Popper, 1971).

Mais, dirait-on, nous ne sommes plus au dix-huitième siècle, ni même au milieu du vingtième. La présence de données tellement massives qu'elles seraient totalisantes permet sûrement de revisiter cette objection, et de poser l'induction comme principe épistémologique fondamental, rendant obsolètes la formulation d'hypothèses et leur vérification par la méthode expérimentale. Cet argument est implicite dans les études qui avancent une méthode *data-driven*, par opposition à *theory-driven*, considérée comme inférieure. Il trouve un appui dans le succès des GAFAs et leurs algorithmes basant leurs prédictions sur les données d'observations passées, sans hypothèse a priori. Il est formulé explicitement par certains auteurs qui affirment par exemple : « Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. (...) There is now a better way. Petabytes allow us to say: "Correlation is enough." (...) Correlation supersedes causation » (C. Anderson, 2008). Il est même contemplé par les physiciens du plus grand accélérateur de particules au monde, aux prises avec des hypothèses qu'ils n'arrivent pas à confirmer, réfuter ou étudier (ATLAS Collaboration, 2018).

Si on envisage une refonte des sciences naturelles, et de toute science à en croire Anderson, il faudrait au moins avoir une réflexion épistémologique sérieuse qui se base sur des propositions justes. Or les arguments concernant la meilleure fiabilité des données numériques, l'auto-suffisance de la corrélation, la supériorité d'une approche fondée sur les données, reposent sur une perception dénaturée des données massives et de leurs outils d'analyse, une « mythologie » selon Boyd et Crawford (Boyd & Crawford, 2012). Pour commencer, il est faux de penser que les données seraient un rendu direct, sans filtre, de la réalité « objective ». Cela semble être une évidence, dit comme cela, mais la facilité avec laquelle le glissement implicite est fait, oblige à l'énoncer clairement. Ce glissement fait l'ellipse de l'heuristique, escamotant les techniques, méthodes et technologies nécessaires à la collecte de ces données. Une IRM cérébrale n'est pas un cerveau : c'est un traitement d'un signal électromagnétique lié aux atomes d'hydrogène des molécules constituant les cellules et les espaces extracellulaires formant le cerveau, en un signal électrique, puis numérique, puis lumineux, que j'interprète ensuite (par une

série d'autres traitements du signal) comme représentant le cerveau. Il y a des niveaux multiples de traitement du signal et d'abstraction qui font qu'on se retrouve avec une donnée qui se décale de l'objet qu'elle représente. Ce traitement peut être vu comme une traduction, une interprétation, avec des pertes possibles d'information à chaque niveau, et des perturbations du signal par du bruit (Berry, 2011).

En plus d'être une interprétation de la réalité, les données en sont une interprétation partielle et fragmentée, malgré les ambitions totalisantes des différents acteurs du Big Data. Toute donnée, qu'elle soit numérique ou non, dépend des outils utilisés pour la recueillir, l'enregistrer, la garder en mémoire. Ces outils ont des caractéristiques techniques, des capacités, des limitations, des possibilités d'utilisation, qui sont conditionnées par le contexte technologique, économique, social et culturel de leur conception et de leur utilisation. Pour illustrer, l'écriture n'était pas accessible à tout le monde à travers les siècles, très restreinte au début, puis s'ouvrant peu à peu, excluant certains groupes sociaux à certains moments, sous l'influence de la maniabilité des technologies d'écriture et de considérations socio-culturelles, et conditionnant ainsi le type de données qui seraient archivées et les sources dont elles proviennent. De même, la limitation du nombre de caractères dans les SMS (Short Messaging Service) était due à des contraintes de bande-passante liées à la structure du réseau de téléphonie mobile utilisé (Taylor & Vincent, 2005), contrainte technique. La limitation du nombre de caractères dans les messages Twitter, bien que basée sur la limitation précédente, est un choix plus commercial, surtout maintenant. Les déterminants techniques et sociaux modulent les types d'informations encodées, la manière dont elles sont encodées, les groupes sociaux dont elles proviennent etc, créant ainsi des biais, parfois évidents, parfois plus sournois, dans le corpus de données. Or tenir compte de ce contexte quand on construit un algorithme n'est pas chose aisée, si elle n'est pas complètement ignorée. On peut en voir les conséquences dans les biais raciaux et sociaux qui apparaissent dans certains algorithmes (Buolamwini & Gebru, 2018; « Data Harm Record », 2017; Zou & Schiebinger, 2018). L'idée que les données numériques seraient représentatives de la diversité d'expériences et d'observations humaines n'est pas fondée. En psychiatrie, où l'on est aux prises avec la marginalisation sociale du point de vue sociologique, et l'aliénation sociale déshumanisante (Oury, 2006a), ceci oblige à même davantage de précaution par rapport à la représentation que font les données numériques des malades.

La fiabilité des données massives, au sens de l'absence d'erreurs dans les données, et la fiabilité de leur mise à disposition, est un autre point à considérer. La

numérisation a augmenté la vitesse et le volume de recueil des données, mais n'a pas éliminé le risque d'erreur. Que ce soit au moment de la saisie manuelle, ou dans les processus d'extraction, de traitement/transformation ou d'archivage, des erreurs ou des omissions peuvent infiltrer toute base de données, problème qui s'amplifie lorsque les bases deviennent de plus en plus massives ou sont entrecoupées pour une analyse, (boyd & Crawford, 2012; Chervenak et al., 2000; Khare et al., 2018, 2017). Peu de résultats de contrôles qualité des données massives en médecine sont publiés, mais des efforts importants de surveillance et de conception de l'architecture des bases paraissent nécessaires pour avoir une qualité satisfaisante, efforts qui ne garantiront toutefois pas l'absence d'erreur ou la possibilité de gérer toutes les erreurs identifiées (Chervenak et al., 2000; Khare et al., 2018, 2017).

Tous ces arguments mènent à un constat : les données numériques étant médiées par un dispositif numérique, elles ne s'offrent pas à nous directement, pas dans le sens d'une note manuscrite par exemple, qui s'offre à notre vue directement après avoir été écrite, sans nécessité d'une médiation supplémentaire. En pratique, à moins d'être un certain personnage de film de science-fiction¹⁹, nous n'accédons pas directement, dans l'usage, aux octets qui constituent cette donnée. Et inversement, pour qu'un objet ou un phénomène soit traduit en une donnée numérique, il doit littéralement être analysé, c'est-à-dire, « dé-composé », découpé, séparé en éléments discrets. C'est la traduction du continu en discret qui est au cœur de la numérisation : « it is the *translation of the continuum into the discrete* that marks the *condition of possibility for computability*²⁰ (...) computation is limited to specific temporal durations and symbolic sets of discrete data to represent reality » (souligné par moi, Berry, 2011). On remarquera, en fait, que ce principe organisateur s'étend au-delà du domaine informatique. À titre d'exemple, on peut observer l'évolution de la conceptualisation américaine de la maladie mentale, allant d'une conception analogique (dans le DSM-2 par exemple, avec sa nomenclature purement descriptive), à une logique algorithmique (DSM-3 et 4, et dans une large mesure DSM-5), à l'organisation « computationnelle » du système RDoC qui est littéralement une matrice (voir Annexe 2). Il semblerait que la condition de calculabilité ou de « calcul-ité » inhérente aux données massives, influence non seulement les modes de saisie des données, mais

19 Neo, dans le film *The Matrix* (The Wachowski Brothers, 1999).

20 « Computability » ne se traduirait pas tout à fait par le terme « numérisation », dans le seul sens d'encodage. Il fait référence à « computation » (calcul, en Français), domaine qui examine ce qui peut et ne peut pas être calculé (Berry, 2011).

également la manière dont on pense nos catégories de la réalité, notre propre ontologie de domaine en quelque sorte.

Il sera peut-être plus aisé de comprendre la place des données numériques entre objectivité et subjectivité, si l'on introduit le concept d'artificialité tel qu'il a été développé par Herbert Simon (Ganascia, 2010). En effet, les données numériques ne sont ni des objets de la nature, ni des sujets évidemment, mais appartiennent à ce domaine particulier des systèmes artificiels définis de la sorte (Ganascia, 2010) :

- ils sont le produit des activités d'êtres intelligents
- ils sont une imitation approximative de la nature
- ils peuvent être caractérisés selon leurs fonctions, leurs buts et leurs capacités d'adaptation
- ils peuvent être abordés selon des termes descriptifs (une analyse structurelle, architecturale, physique) ou impératifs (analyse des fonctions et des buts pour lesquels ils étaient produits).

On peut voir alors, qu'en tant que produits d'une intentionnalité subjective, et obéissant à des impératifs subjectifs et non naturels, les données numériques ne sont pas de purs objets, comme les objets naturels. Dans ce contexte, l'idée d'une opposition entre la recherche guidée par les données (*data-driven*), et celle guidée par hypothèse (*theory-driven*), est étrange. Elle l'est d'autant plus, quand on considère que l'inclusion de différentes modalités de données dans une étude découle d'un choix fait par les investigateurs, basé sur un certain modèle hypothétique a priori. Une étude sur des patients schizophrènes qui combine des données d'IRM de diffusion avec des données génomiques et cliniques, procède d'une hypothèse de départ différente d'une étude qui examine le transcriptome et le protéome périphérique, l'une ancrant sa conception de la maladie dans une hypothèse de connectivité cérébrale et de facteurs génétiques, l'autre ouvrant la possibilité à une hypothèse plus systémique, inflammatoire par exemple. Faisons l'argument par l'absurde : l'IRM du genou ne semble pas être un choix de données considéré comme pertinent par les chercheurs en schizophrénie²¹.

D'un point de vue pratique en psychiatrie, on aura vu tout au long du chapitre II, que les volumes de données utilisées dépassent encore rarement les dizaines de milliers

21 Par curiosité, une recherche PubMed de « knee MRI schizophrenia » retrouve deux entrées : une fausse alerte (genou du corps calleux), et un cas clinique rhumatologique chez un patient qui se trouve aussi être schizophrène. Cette preuve de mon argument absurde n'aurait, bien entendu, pas été possible sans l'utilisation de technologies du Big Data.

de sujets, et se situent plus fréquemment dans l'ordre des quelques dizaines et quelques centaines. Le lien au Big Data se trouve davantage, pour la majorité des études, dans la prise en compte de données multi-modales et les situations à haute-dimension, et leur analyse par des outils d'apprentissage automatique. Les modalités explorant une dimension physiopathologique (IRM, EEG, marqueurs sériques, transcriptome) sont plus difficiles à recueillir en masse, que les données cliniques ou administratives. Il y a cependant, comme le recensent Woo et al. pour l'imagerie (Woo et al., 2017), des consortiums qui s'organisent pour faciliter la mise en commun de données recueillies sur différents sites, ce qui ne pourra qu'améliorer la validation des résultats obtenus à partir de ces groupements de recherche. Certaines pathologies posent des défis particuliers, par leur faible prévalence, le mode et l'âge d'apparition des troubles, ou la difficulté à capturer cette population dans le cadre de soins psychiatriques : on pense notamment aux états à haut-risque de psychose, à l'autisme et au suicide. Plus fondamentalement, la difficulté à capturer le phénomène de maladie mentale, c'est-à-dire l'expérience-même du patient, signifie que la plupart des études reposent sur des questionnaires cliniques, en auto-évaluation ou recueillis par un clinicien, dont la validité repose sur les catégories diagnostiques qui sont actuellement critiquées. On peut se demander si cela n'entraînera pas une sorte de circularité dans les recherches freinant l'avancée vers de nouvelles conceptions. Cette interrogation est présente dans les revues de la littérature (Woo et al., 2017), et certaines études tentent justement de se décaler de ces questionnaires en identifiant d'autres façons de traduire les symptômes, de numériser le vécu du patient, que ce soit par l'analyse du langage et du récit des patients (Bedi et al., 2015; Cohen & Elvevåg, 2014; Natalia B. Mota et al., 2012), l'utilisation des réseaux sociaux (Guntuku et al., 2017), ou les suivis par smartphone (Ben-Zeev et al., 2017; Dagum, 2018). Ces études sont encore très préliminaires, mais elles sont cruciales pour que les représentations de la maladie ne se restreignent pas à une description extrêmement rigide et limitée, avec le risque de standardiser l'expérience du patient jusqu'à en perdre la signification et finalement l'évacuer. On ne doit pas perdre de vue la promesse de « granularité » des données offerte par le Big Data. Toutefois, même dans ces nouvelles représentations, il faudrait être vigilant à questionner les hiérarchies qui seront inévitablement introduites entre les données, par exemple l'accéléromètre du smartphone qui serait considéré comme ayant plus de valeur que l'analyse en réseau sémantique du discours du patient.

Cela ne veut pas dire qu'on ne peut rien savoir à partir des données massives, bien au contraire. Mais cela demande que l'on définisse les limites, ou plutôt les contours de ce

savoir, sa nature, pour en saisir certaines conditions essentielles, et en relation, comprendre la place du Big Data. Popper en propose une lecture assez simple d'ailleurs : le savoir humain, et notamment le savoir scientifique, est par essence conjectural, hypothétique (Popper, 1971). Si on fait fi de l'hypothèse comme le proposent certains, on se demande alors ce qu'il pourrait rester du savoir, ou ce qu'il resterait d'humain dans le savoir. Les données numériques massives sont bien une source – et de plus en plus, une ressource – importante, mais imparfaite, et peut-être même plus, mal utilisées. Parallèlement, l'importance ou la prise en compte d'un phénomène ne devrait pas dépendre de sa capacité à être numérisé, au risque de le rendre invisible s'il n'est pas numérisable.

Donc pour être réellement bénéfiques à l'avancée de la médecine et de la science, les données massives doivent être contextualisées, leurs caractéristiques bien cernées pour comprendre à quelles questions elles peuvent répondre, et à quelles erreurs d'interprétation elles peuvent mener, comme celles discutées dans les études génétiques citées précédemment (voir sous-chapitre II.2.c., p. 90). Il faut aussi comprendre ce que les outils d'analyse disent de ces données, ce qui implique de bien connaître ces outils, et leurs conditions d'utilisation et de validité, mais aussi d'expliquer dans une certaine mesure comment ils arrivent à une interprétation, d'ouvrir la « boîte noire » des algorithmes peu intelligibles.

Cependant, une nouvelle connaissance n'est d'intérêt que si elle est diffusée et soumise à des lectures critiques plurielles. D'où l'importance des modes de diffusion des connaissances et leur adaptation aux résultats générés par des approches Big Data, afin de rendre compte de ces connaissances de manière intelligible, permettant aux lecteurs d'en saisir à la fois et de manière égale la signification centrale, la démonstration, et les nuances et incertitudes. Fait est de constater que le modèle prédominant de diffusion des nouvelles connaissances scientifiques et médicales est mal adapté à cette tâche actuellement, avec une fragmentation des connaissances, des formats rigides qui correspondent mal à la description des procédures en apprentissage automatique, et une mise en page qui ne permet pas au lecteur de saisir les informations et les nuances facilement. De plus, l'incitation à la publication et à être cités pousse les auteurs à montrer leurs meilleurs résultats, plutôt que ceux qui sont les plus fiables même s'ils sont négatifs, notamment dans l'écriture des abstracts, qui sont le premier abord d'un article et souvent la seule partie lue. Des évolutions positives se mettent en place, néanmoins, avec la multiplication de revues et de dépôts ouverts, ou d'obligations de publications en accès

ouvert, comme décrit au chapitre I.3.c (voir p. 56). Certaines revues offrent en ligne, une mise en page permettant plus facilement au lecteur de visualiser ensemble les résultats graphiques et textuels, ou de prendre connaissance des références en contexte (les sites des publications JAMA, ou PubMed Central ou Cairn.info par exemple). D'autres offrent un format plus adapté à la description des techniques numériques et des bases de données (Coppersmith et al., 2018). De manière plus pertinente au Big Data, des modalités apparaissent permettant de partager directement les différentes étapes d'une étude et les modèles développés, et permettre un accès direct à la visualisation de ce modèle en fonctionnement, et au code de programmation (Perkel, 2017b, 2017a), éléments cruciaux pour permettre à d'autres chercheurs de tester et valider les modèles rapidement.

De même, pour répondre aux préoccupations concernant la fiabilité des bases de données, certains auteurs proposent de créer des « datasheets for datasets », en quelque sorte des manuels de caractéristiques et d'utilisation accompagnant les jeux de données, des méta-données pour les données (Gebru et al., 2018). D'autres chercheurs, enfin, travaillent à rendre les algorithmes plus intelligibles, à retracer leur « raisonnement » d'une certaine manière (Fauw et al., 2018; Ribeiro et al., 2016a; Ribeiro, Singh, & Guestrin, 2016b). Des efforts plus importants devraient être dirigés vers la mise en correspondance de différentes études, de différents résultats pour progressivement construire une vue d'ensemble sur une question. Les revues participatives comme celle de Ching et al. sur l'apprentissage profond sont un premier modèle (Ching et al., 2018), mais il peut y en avoir d'autres. Le fil directeur ici est l'idée d'une re-contextualisation d'éléments décontextualisés, une réinscription d'éléments discrets, discontinus, dans une sorte de continuité, ou plutôt, pour employer une métaphore très usitée en psychiatrie, de retisser une toile à partir de ces fils séparés.

En érigeant une version dénaturée, mythique, du Big Data comme moyen pour faire de la psychiatrie une discipline aux bases principalement, voire exclusivement, scientifiques, on se perdra : on n'aura ni la rigueur critique d'une science hypothético-déductive, ni l'hétérogénéité créative de lectures proposées par les « sciences de la culture », ni même la possibilité d'examiner ce que les données massives et l'apprentissage statistique peuvent nous apprendre sur l'incertitude, la nature conjecturale du savoir, la coloration de l'épistémologie par la technique, car la vision idéaliste du Big Data ne laisse pas de place aux interrogations, aux nuances et aux incertitudes. On obtiendra peut-être, tout au plus, une efficacité technocratique circulaire, ce qui n'est pas

du tout la même chose que de mieux soigner et s'occuper des malades, ou de mieux faire de la recherche scientifique.

Ainsi, une position de simple consommateur par rapport aux technologies numériques et aux données massives risque de mener à une subversion non seulement du savoir, mais aussi de la pensée, et par extension, de l'autonomie et de la liberté. Plusieurs auteurs s'interrogent en effet sur la question de l'expérience et de la responsabilité médicale à l'heure du développement d'algorithmes d'aide à la décision clinique (Cabitza, Rasoini, & Gensini, 2017; Char, Shah, & Magnus, 2018). Les algorithmes pouvant être *perçus* comme plus rationnels, un médecin se sentira-t-il libre de suivre son propre avis clinique s'il va contre celui de l'algorithme ? Comment cette perte de confiance en son expertise affectera-t-elle le sens clinique du médecin, l'importance qu'il accorde à son expérience clinique, et donc l'importance qu'il accorde à son quotidien de travail auprès des malades ? Inversement, si un algorithme se trompe, comment déterminer la responsabilité de cette erreur ? Ces auteurs écrivent notamment : « The collective medical mind is becoming the combination of published literature and the data captured in health care systems, as opposed to individual clinical experience. » (Char et al., 2018)

En pratique clinique, comme en recherche, on assisterait à une sorte de mise à l'écart, voire un dessaisissement, de la pensée, de la raison humaine, tendance qui, sans ironie de ma part, semble être déjà ancrée depuis un certain temps dans le domaine politique. Plutôt qu'aux délibérations imparfaites d'une collectivité citoyenne, il est fait référence à des « forces » externes, décontextualisées au point d'acquérir un pouvoir mythique : le marché libre, l'Europe (Bourdieu & Wacquant, 2000), et maintenant les algorithmes, l'intelligence artificielle (Rouvroy & Berns, 2010). La référence au mythe est d'autant plus pertinente dans le cas du Big Data, car les pouvoirs de prédiction qu'on lui impute rappellent les oracles des dieux d'Olympe. Même l'idée d'un regard humain sur tout algorithme décisionnel ferait penser à la place des prêtres et des devins servant d'intermédiaires et d'interprètes. À défaut d'avoir su cultiver un autre de nos héritages helléniques, serait-on en train de se rabattre sur celui-ci à sa place ?

Il y a un certain confort, une complaisance, à externaliser le pouvoir décisionnel effectif en dehors de notre raison. Notre responsabilité semble moins engagée, nous aurions moins besoin de nous justifier, de justifier notre raisonnement car il n'y aurait pas de raisonnement à justifier. Cela paraît moins risqué, car indépendant de variations individuelles si difficiles à maîtriser. Face à l'aversion au risque et à l'incertitude, les

référentiels sont rassurants²². Quoi de mieux donc qu'un instrument en l'apparence purement logique ? En renonçant à cette responsabilité, cependant, nous diminuons notre autonomie, au sens kantien de sujet libre. Or dans un lieu de soin, quelle sera la conséquence de cette limitation de liberté ? Sur ce point, le mouvement de la psychothérapie institutionnelle a bien décrit l'influence de l'ambiance, des « entours » sur le soin psychiatrique (Delion, 2014; Oury, 2007a). Que ce soit pour éviter les cloisonnements selon les statuts de chacun, qu'induisent les structures hiérarchiques (structures qui évitent par ailleurs d'avoir à penser) ; ou le glissement vers des « relations dyadiques » qui ne sont plus thérapeutiques mais profondément aliénantes ; ou tous ces différents éléments de fonctionnement institutionnel qui, s'ils ne sont pas interrogés, sédimentent et induisent une toxicité institutionnelle qui entrave le soin ; l'introduction d'une certaine liberté permet de créer des ouvertures sur des « espaces du dire » (Oury, 2006b). La liberté est d'abord liberté de circulation pour les malades dans tous les espaces d'une institution, inscription d'une liberté dans l'espace pour ouvrir vers une liberté de la parole et de la pensée. La liberté, ensuite, est autonomie : pas une autonomie fonctionnelle visant à savoir faire les choses seul, sans aide extérieure, mais la liberté de prendre des responsabilités dans un collectif, pour la co-gestion de certains dispositifs de l'institution de soins (les « clubs thérapeutiques »). L'un des risques d'une érosion de cette liberté serait le retour à un fonctionnement asilaire (Delion, 2014), système fermé profondément aliénant.

Cette « sphéricité », cette fermeture sur soi, se retrouve au niveau épistémologique. On pourrait dire que l'hypothèse introduit une incertitude, c'est une demande, qui suit une logique de « coupure ouverte » selon Lacan cité par Oury (Oury, 2005). Elle contient une possibilité créatrice, une ouverture. Pourrait-on dire même qu'elle contient quelque chose du désir ? L'induction, elle, peut se visualiser comme un système fermé, qui s'auto-référence : comment sais-tu que ce cygne sera blanc ? Parce que tous ceux avant lui étaient blancs. Pourquoi est-ce que ce cygne est blanc ? Parce que tous ceux avant lui étaient blancs. Cette caricature en est à peine une, car le problème dans l'utilisation d'un système inductif comme système prédictif, inférentiel et décisionnel, c'est de créer des boucles de régression à l'infini sur des prophéties auto-réalisatrices. Des exemples

22 Ceci n'est pas une critique des référentiels et des recommandations, qui sont des repères indispensables pour l'organisation d'un système de santé, et des travaux de synthèse et de consensus, jamais parfaits, mais toujours importants. Ce qui est remis en question est leur instrumentalisation, la manière dont ils peuvent être utilisés dans des logiques qui ne sont pas toujours celles du soin.

concrets de ces effets ont été documentés, notamment en ce qui concerne les outils de police prédictive (« Data Harm Record », 2017).

On voit qu'un tel système n'a pas de sens, au sens figuré comme au sens propre. Et justement, il n'est pas nécessaire qu'il en ait un pour être efficace la plupart du temps. Ce n'est pas la recherche de sens qui est au cœur des algorithmes prédictifs ayant eu le plus de succès jusqu'ici, mais une recherche de performance, d'efficacité. Google Translate, par exemple, ne traduit pas en « comprenant » le texte, mais en établissant des corrélations à travers la comparaison de larges corpus de textes écrits dans les deux langues à traduire. Google Translate ne pourra pas générer une fiche de lecture comme celle demandée aux écoliers après la lecture d'un livre. Dans un autre domaine, pour les algorithmes de classification d'images, des « attaques adversaires » (adversarial attacks), ou des perturbations de l'image qui seraient imperceptibles ou peu perturbatrices pour un humain, peuvent mettre les algorithmes en grande difficulté menant à des classifications très étranges (Rosenfeld, Zemel, & Tsotsos, 2018). Ceci fait penser que les processus sous-tendant la performance des algorithmes sont bien différents, étrangers même, de ceux qui permettent aux humains de comprendre les choses, d'y mettre du sens. Dans une étude très sérieuse, mais dont le titre peut prêter au rire, « Could a Neuroscientist Understand a Microprocessor? », Jonas et Kording essaient de démontrer que les méthodes d'analyses de données communément utilisées en neurosciences ne permettent pas d'accéder à une réelle compréhension du fonctionnement d'un microprocesseur, système pourtant bien plus simple que le système nerveux (Jonas & Kording, 2017). Ils réfutent ainsi l'idée selon laquelle ce serait un manque de données qui limiterait actuellement notre compréhension du cerveau, et que l'accumulation de données massives résoudrait ce problème.

Resituer la conception des données massives en objet artificiel permet de redéfinir leur cadre épistémologique et éviter les mouvements de rejet ou d'idéalisation. Jean-Gabriel Ganascia propose en effet, en discutant l'épistémologie de l'intelligence artificielle, de la définir en « domaine intermédiaire », ne pouvant être réduite à une science de la nature, une science théorique ou une science de la culture, et ainsi dérivant ses méthodes et ses objets d'étude des trois (Ganascia, 2010). Ce raisonnement peut aisément être étendu aux données massives. Et ne serait-ce pas aussi le propre de la psychiatrie d'être un « domaine intermédiaire » ?

Ce que le Big Data pourrait offrir de plus précieux, donc, ce sont peut-être des points de départ pour de nouvelles hypothèses, plutôt qu'un « savoir » pré-mâché ou des

décisions autoritaires. Avec des méthodes qui gèrent et décrivent l'incertitude, le Big Data pourrait nous aider à mieux identifier les problématiques, à poser des questions scientifiques et humaines plus pertinentes : à montrer ce que l'on ne sait pas. Ses outils de visualisation et d'organisation des données pourraient ouvrir vers des modes de transmission des connaissances qui soient plus intelligibles. L'élargissement de l'accès à l'information, et de la création de données, à des publics qui en étaient jusqu'ici exclus crée des tensions mais également des ouvertures, de nouvelles possibilités d'existence, qu'il conviendra de naviguer avec soin.

3. Être atomisé

Cet élargissement « ubiquitaire » (Pucheu, 2014) de la médiation numérique interroge non seulement la pratique, quelle est notre relation à ce que l'on fait, ou la pensée, comment appréhende-t-on le monde autour de nous, mais aussi l'identité, quelle est notre relation à l'Autre, et par là comment construit-on son Soi ? Difficile de procéder à une synthèse des faits sur cette question, quand la génération des *digital natives*, ceux ayant grandi avec la présence du numérique à portée de main, et pas seulement avec le *personal computer* familial dans un coin du salon, est encore en développement. La question ici n'est pas celle des pathologies associées ou ayant à leur centre l'utilisation des outils numériques. Sur ce sujet, en effet, des données commencent à apparaître, conduisant par exemple à l'inclusion dans le DSM-5 de l'*Internet Gaming Disorder* comme trouble nécessitant des études supplémentaires. Certaines études documentent les effets délétères d'un usage excessif du smartphone ou d'internet sur l'humeur, l'anxiété et les relations sociales (Kim, Jang, Lee, Lee, & Kim, 2018; Kuss, Harkin, Kanjo, & Billieux, 2018). Il y a probablement des caractéristiques propres aux technologies numériques, notamment dans leurs interfaces, qui pourraient, en interagissant avec des vulnérabilités propres à chaque individu, mener à des processus pathologiques, ou constituer une réponse mal adaptée à ces vulnérabilités.

Mais la question que nous allons essayer de traiter sera plutôt la suivante : est-ce que la numérisation ubiquitaire change fondamentalement la subjectivation et la subjectivité, la construction et la reconnaissance du Soi, d'une identité, ou est-ce qu'elle vient seulement se confronter, que ce soit de manière bénéfique, destructrice ou neutre, à un Soi qui se construit indépendamment d'elle ? Il n'est probablement pas possible, en

l'état actuel des connaissances et de notre expérience, de pouvoir répondre de manière satisfaisante à cette question. Toutefois, des éléments de réflexion peuvent être proposés, et notamment mis en lien avec le soin en psychiatrie.

Pour une description plus phénoménologique du vécu des patients et des soignants par rapport à l'utilisation des technologies numériques dans le soin, il faut se tourner vers des études ayant une approche sociologique et anthropologique, la littérature médicale et scientifique explorant cette question plutôt d'un point de vue opérationnel, instrumental (faisabilité, acceptation, rejet par exemple). Une revue de la littérature critique sur ce sujet, abordant des applications très diverses du numérique en santé à travers une grande variété de publications, souligne toute l'ambiguïté de ces technologies pour les patients et les soignants (Lupton, 2014a). Allant des plateformes d'échanges d'informations médicales (entre patients, soignants ou entre soignants et patients), à la télé-médecine, à l'intégration du numérique aux politiques de santé, à « l'incarnation numérisée » (« digitised embodiment »), Deborah Lupton relève les interrogations soulevées par l'observation de ces expériences d'utilisation du numérique en santé :

- la tension entre l'utilisation de plateformes numériques qui ouvrent de nouvelles possibilités de parole, d'expertise et d'échanges collectifs, et le risque de « commodification » de la parole du malade (« commodification of patient opinion », Lupton, 2014b) à travers ces plateformes, par des utilisations commerciales parfois même à l'insu des personnes,
- l'ambivalence de la télé-médecine, du suivi à distance et des mesures en continu, perçus comme libérateurs ou au contraire comme forme de surveillance invasive et de contrôle (au sens foucauldien), des dispositifs qui sont soit « bricolés » par les soignants soit s'imposant et contraignant leur pratique,
- le discours politique sur l'engagement et l'autonomie des patients soutenus par le numérique, la santé passant progressivement d'une responsabilité collective à une responsabilité individuelle, avec le risque de désinvestissement du collectif, allant de pair avec certaines conceptions du néo-libéralisme politique,
- la numérisation de plus en plus complète de tous les aspects et recoins du corps humain introduisant de nouvelles formes de compréhension mais aussi une nouvelle approche, tant conceptuelle que physique, du corps en médecine.

Ce « digitised embodiment », la numérisation de l'incarnation, du corps vécu, du *Leib* – corps en Allemand, dans le sens utilisé par Gisela Pankow dans son travail avec les patients schizophrènes (Pankow, 2009) – peut nous donner un point d'accroche pour

penser la construction du Soi à l'ère du numérique. Avec la miniaturisation de différents types de capteurs, leur insertion dans des objets largement disséminés et utilisés, voire indispensables, comme nos smartphones, la facilité avec laquelle on peut archiver et partager une trace visuelle (photographie ou vidéo), la numérisation en continu de notre corps tant dans son anatomie statique que dans son fonctionnement dynamique devient de plus en plus possible et même désirable. En témoignent l'adoption dans un cadre non-médical des appareils et des applications permettant le suivi et l'archivage de données aussi diverses que la quantité de mouvements effectués, de calories consommées, de distance parcourue, les signes associés au cycle menstruel, ainsi que la dissémination et le partage de descriptions détaillées de symptômes, assorties d'images et de vidéos sur les réseaux sociaux (Lupton, 2014b, 2014a; « Technology - Fitbit », s. d.).

Ce *Leib* numérisé est traduit en d'innombrables points de données (data points), qui par la logique-même de *computationality*, transforment le continu en discret, en parties séparées, indépendantes, qui tentent de quantifier et de mesurer le vécu. Cette quantification du corps ne permettrait pas seulement de suivre quelques mesures physiologiques importantes dans le cadre d'une maladie (la glycémie chez un patient diabétique, la pression artérielle chez un hypertendu), encore qu'il faille bien juger l'indication de mesures continues chez des personnes stables par ailleurs. Non, cette mouvance vers la numérisation totale du corps, qui s'érige même en mouvement dans des communautés de « self-trackers » (« Quantified Self - Self Knowledge Through Numbers », s. d.), tend vers la construction d'un « Soi quantifié », promettant un accès à la connaissance de soi à travers une quantification minutieuse des fonctions corporelles et cognitives (Lupton, 2016; Wolf, 2010). Mais ce faisant, on procède à une externalisation de ces bouts d'être, à leur réification, comme des objets jetés en dehors du corps et qui prennent une existence propre. Une externalisation qui fige aussi, posant un signifiant sur quelque chose qui était jusque-là sous-jacent.

Quel est l'effet sur la subjectivité d'une telle représentation désincarnée et fragmentée du vécu, amenant sur le devant de la scène des fonctions corporelles qui en temps normal, ont une autonomie propre et sollicitent notre attention seulement en cas de menace ? Comment ces mesures externalisées prennent-elles sens ? Nous appartiennent-elles même encore ? Comment modifient-elles notre regard et notre lecture de notre corps : source d'inquiétude, voire d'angoisse, ou de réassurance ; d'expertise ou de perte de contrôle ; une attention mieux dirigée ou submergée ; collage à cette représentation, évacuation du corps de la subjectivité, ou intégration harmonieuse ? Une

analogie imparfaite serait ici intéressante. Il existe un domaine de la médecine où une surveillance continue et simultanée de nombreuses fonctions corporelles a lieu, et est même absolument nécessaire : dans le cadre des soins d'anesthésie-réanimation, les soins « intensifs », où la dépendance, l'instabilité ou la perte d'autonomie du corps est tellement extrême, « critique », qu'une externalisation des fonctions vitales est nécessaire pour le maintien de la vie. Assisterait-on là à un état critique généralisé ?

Si la numérisation du vécu se contentait de nous offrir une représentation, certes fragmentée, mais une parmi d'autres à pied égal, ce ne serait peut-être pas si problématique. Seulement, là ne s'arrête pas le projet d'une numérisation « ubiquitaire », tel qu'il semble se déployer (Pucheu, 2014; Weiser, 1991). De plus en plus de nos vécus sont médiés par des objets informatiques, créant des couches d'abstraction de plus en plus complexes et difficilement accessibles, entre nos expériences physiques quotidiennes et les conséquences physiques qu'elles ont sur le monde. Ce découplage, quand il fonctionne bien, doit justement se faire sans que l'on s'en rende compte, doit « laisser les ordinateurs disparaître en arrière-fond » (« allows the computers themselves to vanish into the background », (Weiser, 1991), comme lorsqu'on conduit une voiture à moteur à injection électronique. Son but est d'être invisible, de donner une illusion de continuité, de fluidité (*seamlessness* en Anglais, littéralement « sans couture ») entre l'action humaine et la conséquence sur le monde physique, illusion qui repose en grande partie sur la différence entre la temporalité humaine et algorithmique (Berry, 2011). Même au-delà de ça, elle « promet de liquider les incertitudes et de rectifier l'incomplétude qui préside à nos rapports protéiformes avec le monde » (Pucheu, 2014), « invisibly enhancing the world that already exists » (« accentuant ou améliorant de manière invisible le monde qui existe déjà », Weiser, 1991).

Donc le numérique ubiquitaire ne serait pas seulement une médiation de nos représentations du monde, mais aussi de notre vécu du monde, jusque dans notre chair même, notre *Leib*. On a vu qu'il s'agissait d'une médiation fragmentée, qui fracture le continu en éléments discrets calculables. Cette médiation pourrait aussi transformer notre inscription au sein de deux dimensions qui nous structurent pourtant fondamentalement : l'espace et le temps. En effet, l'algorithme et les données ont une spatialité qui peut être décrite comme élastique. Dispersée à travers le monde dans des réseaux complexes qui ne s'offrent pas à la vue de l'utilisateur simple, et dont la topologie logicielle n'est pas superposable à leur topologie géographique, cette spatialité vient se télescoper dans le contexte spatial plus circonscrit de l'utilisateur, et se trouve limitée dans la structure

physique de la mémoire de l'ordinateur (Berry, 2011). Elle en vient même à se désolidariser de la terre ferme, flottant dans un « cloud » (nuage). À l'être humain qui l'utilise, le numérique fait vivre une contraction extrême des distances par ce qu'il offre à portée de main, tout en creusant davantage d'autres distances en diminuant l'importance de son environnement physique proche, reconfigurant ainsi la spatialité de l'être humain à l'image du numérique. On achète ses livres sur Amazon, ou mieux encore, on les télécharge, plutôt que de visiter la librairie du coin ; on se sent plus d'affinité avec des amis à l'autre bout du monde avec qui on peut maintenir un lien quotidien, alors on peut et on veut moins investir et s'investir dans ce nouveau quartier où on vient d'emménager.

Le temps aussi paraît étranger sous l'angle de la temporalité numérique. Il semble disparaître avec des données qui sont à jamais enregistrées dans des mémoires qui n'oublient pas. Il est vastement accéléré dans la vitesse des microprocesseurs et des moyens de transmission, par rapport à la temporalité de l'action humaine (Berry, 2011). Cette vitesse-là est cachée des humains, se retirant en arrière-plan pour soutenir l'impression de fluidité des actions numériques. En revanche, la transformation progressive de l'Internet, d'une architecture de pages à contenu statique, à des flux continus d'informations en temps-réel (Berry, 2011), plonge l'utilisateur au sein d'un débit d'information tellement massif et rapide, qu'il le maintiendrait dans un présent permanent, incapable de se relier au passé, ni d'élaborer un futur²³. Le temps disparaît ou reste collé, perdant son sens.

Dans ce paysage spatio-temporel altéré, l'être humain n'est pas seulement un utilisateur, un consommateur de données en interaction avec un outil informatique, il les produit aussi activement et passivement, en sa qualité de « prosumer », contraction de « producer » et « consumer », producteur et consommateur (Lupton, 2014b). Or l'agencement algorithmique prédominant actuellement autour de ces données, avec la qualité réflexive de sa logique inductive, tend à refermer cette boucle sur elle-même, dans une sorte d'auto-référencement extrêmement appauvri et homogénéisé, car bien qu'il ne repose pas sur les mêmes données pour chaque individu, il repose sur une même classe de processus algorithmiques. En effet, Berry écrit : « Indeed, this may even point to a homogeneity of being in the digital 'age', as we become a being whose existence is

23 Un lien très intéressant peut être fait ici avec la philosophie de Martin Heidegger, notamment autour de sa conception de la temporalité humaine. Cette réflexion ne sera pas abordée ici, mais elle est développée par plusieurs abords dans l'ouvrage de David M. Berry, *The philosophy of software: code and mediation in the digital age* (Berry, 2011).

mediated by identical computational processes » (Berry, 2011). Ce qui pourrait se constituer donc, c'est une homogénéité repliée sur soi, sans confrontation à une altérité structurante, sans possibilité de rencontre avec d'autres subjectivités, débarassées de leur « ipséité » (ce qui en fait des sujets propres) et réduites à leur « mêmeté » (Pucheu, 2014; Ricœur, 1988).

Rappelons-nous que le numérique ubiquitaire promettrait de « rectifier l'incomplétude » de nos rapports avec le monde (Pucheu, 2014). Là l'objectif de *seamlessness*, de fluidité, peut se comprendre non seulement comme la disparition de la médiation donnant l'impression d'une continuité de l'*action* humaine sur le monde, mais comme la continuité entre l'*individu* et le monde. Autrement dit, la frontière entre le Soi et l'extérieur n'existerait pas, le monde extérieur étant une extension du Soi, tout étant Soi. Seulement, l'inverse de cette proposition serait que le Soi n'existe pas... Une autre lecture de cette *seamlessness* au monde rejoindrait les questions que nous avons abordées autour de l'objectivité des données. Le mythe d'un collage parfait des données au monde, ne serait-ce pas d'une certaine manière, le souhait d'une adhésion totale du signifiant au signifié, d'une désymbolisation du monde ?

Toutes ces représentations, ces médiations, ces illusions reposent sur une base extrêmement fragile, « brittle » en Anglais, cassant, désignant quelque chose d'apparence solide, mais qui peut facilement se fracasser (Berry, 2011). Que ce soit par la fragilité du matériel informatique et sa sensibilité aux aléas physiques, les limitations physiques posées par ce matériel, la complexité du *code* numérique avec les *bugs* qui peuvent s'y introduire, les conflits entre fonctions d'un logiciel ou entre logiciels, les incompatibilités de mises à jour, la vulnérabilité de la mémoire à la corruption, du système entier aux cyberattaques, toute cette infrastructure s'adapte mal à l'inhabituel, à la surprise. Une subjectivité construite sur cette base en hériterait-elle les fragilités et le manque de résilience ? Ces imperfections nous montrent aussi que l'idéal de fluidité, de *seamlessness*, n'est bien sûr pas encore réalisé, s'il le sera un jour. Dans ce cas, toutes ces petites perturbations offrent-elles des ouvertures pour qu'une subjectivité propre échappe à cette médiation ? Ou ne sont-elles que des parasites (Berry, 2011) obligeant à leur prêter attention, forme de contrôle de l'outil technologique sur soi ?

Certaines œuvres de culture récentes capturent finement ces interrogations humaines autour de l'ubiquité numérique, comme la série télévisée américaine *Mr. Robot* (Esmail, 2015). Par le récit psychopathologique du protagoniste, ingénieur informatique de génie et aussi *hacker*, inscrit dans une intrigue politique, économique et technologique

complexe, la série aborde les questions de subjectivité, de rapport au temps, de rapport à l'Autre, soulignant, plus par les choix cinématographiques que par le dialogue ou l'intrigue, l'aliénation des personnages, leur déconnexion, toutes ces rencontres manquées. La série *Black Mirror* (Brooker, 2011), quant à elle, à travers des vignettes d'une grande originalité, souvent humoristiques ou frisant l'absurde, suscite un sentiment d'étrangeté troublante face aux nouvelles technologies, cette « Unheimlichkeit », ou « the uncanny », comme a essayé de le décrire Freud (Freud, Strachey, Cixous, & Denomé, 1976) en s'appuyant sur la littérature fantastique.

Pour nous aider à naviguer ces différents faisceaux d'idées, il serait peut-être utile de nous tourner vers la clinique, plus particulièrement vers celle de la psychose, en tant qu'elle interroge aussi la question de l'identité chez l'être humain. Pour être très clair, il ne s'agit absolument pas ici de faire des hypothèses psychopathologiques qui seraient sans fondement, ni de suggérer que le numérique serait « schizophrénogène », ni de faire de fausses équivalences entre une sémiologie source de souffrance humaine réelle, et des idées encore hésitantes sur les effets du numérique. De telles spéculations n'ont pas leur place ici. Cet exercice doit se voir presque comme un exercice littéraire, si on peut se le permettre, pour ouvrir une fenêtre sur une compréhension de l'être humain, comme lorsque l'emploi d'une métaphore par le poète permet un éclaircissement soudain, la création d'une signification nouvelle.

Avec ces précautions, on peut dire que la clinique de la schizophrénie nous met face à ces questions de discontinuité, d'un tout qui manque de consistance, qui ne tient pas ensemble, ainsi qu'aux questions de porosité entre soi et l'extérieur, les autres. On retrouve des instances de cette discontinuité dans la dissociation, dans les troubles du cours de la pensée par exemple, la diffluence et la discordance. Le travail de Gisela Pankow avec des patients schizophrènes montre aussi cette dissociation au niveau corporel, décrivant des corps vécus comme fragmentés, les morceaux perdant toute connexion entre eux (Pankow, 2009). Cela s'entrevoit dans le discours d'une jeune patiente schizophrène citée par Pankow :

« Je ne peux pas vivre à la fois dans ma tête et dans mon corps. C'est pourquoi je ne réussis pas à être *une seule personne*. Quand je suis dans ma tête, j'oublie aussitôt mon corps. (...) *Je suis la chose que je vis*. Mais ce quelque chose peut vivre ou être seulement un objet. (...) Le plus beau jour de ma vie est le jour où se passa l'évènement suivant : on avait déménagé du mobilier et sous un meuble, je vis se faufiler rapidement

un crabe bleu. *J'eus l'impression que je vivais dans le crabe.* » (souligné par l'auteur, Pankow, 2009)

Ces questions vont au cœur de la problématique schizophrénique, celle de l'identité, du Soi. Le discours de cette patiente montre justement la difficulté à délimiter un Soi séparé du monde extérieur et des autres. Cet aspect se manifeste dans les hallucinations acoustico-verbales par exemple, et dans certaines idées délirantes qui témoignent de l'incertitude d'avoir la maîtrise et la possession de ses pensées : l'automatisme mental, l'écho de la pensée, les idées délirantes d'influence. La déstructuration spatiale empêche une inscription temporelle, phénomène que l'on observe également de façon marquante dans un état délirant rare, non-schizophrénique cette fois, le syndrome de Cotard qui peut s'observer dans la dépression mélancolique : la négation des organes et du corps, leur putréfaction, le délire apparemment paradoxal d'immortalité, non pas au sens d'une vitalité illimitée mais de ne pas pouvoir mourir, de se situer, en somme, hors d'une temporalité et d'une corporalité humaines.

Ces différentes pistes suggèrent des manières dont l'usage actuel du numérique pourrait altérer le rapport à l'Autre et la subjectivité. Pour utiliser une image – encore une, ce serait comme si les technologies numériques telles qu'elles sont actuellement agencées autour de nous et utilisées par nous, créeraient un espace-temps distordu, où tout convergerait vers un Soi massif : en somme un trou noir. Il n'est pas du tout certain que tous ces mécanismes soient en jeu, ni même que l'évolution de nos subjectivités se fasse forcément dans ce sens. Mais, on est en droit d'interroger la direction qui risque d'être suivie. Pour prendre un exemple très proche de notre actualité médicale, le nom du nouveau plan de santé du gouvernement, annoncé en septembre 2018 (Ministère des Solidarités et de la Santé, 2018b), « Ma santé 2022 : un engagement collectif », interpelle. Dénommé presque comme un dossier informatique (« Mes documents », « Mes vidéos »), il se veut, l'on suppose, comme une référence à la médecine personnalisée, ici prioritaire. Mais on se demande de quelle médecine personnalisée il s'agit. Une médecine qui reposerait sur la seule logique computationnelle, fragmentaire, auto-référencée et fragile, renferme le patient dans une boucle aliénante. Une médecine qui insisterait de manière disciplinaire sur la responsabilité personnelle du patient, pendant que les établissements publics se dessaisissent ou sont dessaisis de leurs responsabilités, serait en contradiction avec la volonté affichée d'un « engagement collectif ». À ce titre, cette observation de Gilles Deleuze paraît comme un avertissement : « la nouvelle médecine « sans médecin ni malade » qui dégage des malades potentiels et des sujets à risque, qui ne témoigne

nullement d'un progrès vers l'individuation comme on le dit, mais substitue au corps individuel ou numérique le chiffre d'une matière « dividuelle » à contrôler. » (Deleuze, 2007 cité dans Rouvroy & Berns, 2010)

On sent bien, que face à la problématique d'une subjectivité mise à mal par l'altération du lien à l'Autre, il faudrait peut-être remettre un peu d'Autre. Mais comment devrait-on s'y prendre ? Et qu'est-ce que cela voudrait dire, de remettre de l'Autre ? Revenons au corps, à ce *Leib*, par lequel nous avons commencé cette discussion. Autour du corps, Jean Oury développe, le temps d'un écrit, l'idée des « entours » et de la fonction « scribe » du *Leib* (Oury, 2007a). Dans une institution de soins, l'ambiance, ce qu'Oury nomme les « entours », ce sous-jacent, a autant d'importance dans le soin que les actions manifestes et volontaires, voire même plus, car ces entours constituent ce dans quoi baignent au quotidien soignants, soignés, actes, paroles, gestes. Pour des malades dont l'une des caractéristiques de la maladie est un manque de distinction entre le Soi et le monde, l'attention portée à ces entours, au travail sur les entours est d'autant plus important. Le corps pourrait être justement cette surface, ou cette interface, qui recevrait quelque chose de ces entours, à tout moment, et peut-être même davantage quand l'accès à la parole est encore compliqué. Par ce travail collectif sur les entours, dans le quotidien et dans la relation, et par la « constellation transférentielle » qui apparaît autour du malade, des ouvertures se créent, jusqu'à ce qu'à un moment, une de ces ouvertures qui fait effet de surprise permette une « inscription », quelque chose qui marque le malade, et en retour les soignants, et qui peut commencer à prendre sens pour le malade.

Comment allons-nous donc, malgré, ou en faisant avec, ou en transformant cette fragmentation de tous *nos* vécus de soignants, pouvoir créer les conditions pour que quelque chose s'inscrive chez le malade, dont la fragmentation du vécu est bien réelle et non médiée ? Comment faire pour ouvrir une possibilité au transfert dans cette ère numérique ? C'est-à-dire, comment travailler ces « entours » numériques ?

Conclusion

Revenons aux sources, avant de conclure. Le mot « algorithme » nous vient, via l'Espagne médiévale, de l'Arabe Al Ḥwārizmī, surnom du mathématicien Abdallāh Muhammad ibn Mūsā (Centre National de Ressources Textuelles et Lexicales, s. d.-a). Al Ḥwārizmī veut dire littéralement « celui de Hwārizm », c'est-à-dire celui venant de Khwarezm, ou Chorasmie, un territoire situé en Asie Centrale. Un signifiant profondément inscrit dans un terre et une histoire, pour un signifié souvent perçu comme existant au-delà de ces cadres humains.

Ce travail a essayé de retracer le contexte du Big Data, dans sa technique, son architecture, sa politique et sa culture, avant de procéder à une évaluation méthodologique de sa contribution scientifique au domaine psychiatrique. Cette recontextualisation était nécessaire afin de tenter par la suite une analyse critique de l'impact du Big Data sur la pratique et l'épistémologie médicales et psychiatriques, et sur la subjectivité.

Loin de la vision idéalisée, que partagent, ou que du moins véhiculent, nombre de politiciens, ingénieurs, hommes d'affaires, scientifiques et médecins, la réalité du Big Data est celle d'un objet artificiel de création bien humaine : utile, très ingénieux, mais aussi très fragile, imparfait, insatisfaisant. Et pour s'en rendre compte davantage, il est intéressant de chercher au-delà de ses propres cases et catégories. L'expérience du Big Data en climatologie, domaine pourtant bien loin de la psychiatrie, impliquant des temporalités géologiques difficiles à cerner, arrive cependant à des conclusions semblables (Faghmous & Kumar, 2014) : on ne peut pas faire l'impasse de penser et mettre de côté le sens critique. Dans les deux domaines, malgré la promesse d'une révolution des pratiques par le Big Data, la réalité en est bien loin. Les efforts et les atouts du Big Data en psychiatrie pourraient justement être mieux canalisés, sans cette frénésie de la prédiction, vers des utilisations et des méthodes plus pertinentes.

Face à cette complexité presque écrasante, de la subjectivité humaine, de notre avenir géologique, la sidération peut être un mécanisme de défense, mais qui nous fera vite défaut. Alors devrait-on s'informer ou se former au Big Data, comme préconiseraient certains plans gouvernementaux ? Si cette exhortation renferme une volonté de former des consommateurs efficaces et dociles de cette technologie, il faudrait trouver un moyen

de la subvertir. Car comme beaucoup d'inventions humaines utiles et ingénieuses, le Big Data aura vite fait de limiter notre liberté, voire de nous déshumaniser, si nous en avons une utilisation irréfléchie.

Il faudrait donc pouvoir se parler et réfléchir ensemble à l'introduction, la conception et le bien-fondé de ces technologies en psychiatrie : malades, ingénieurs, psychiatres, infirmiers, éducateurs etc. Il faudrait cultiver, ce que Alan Turing appelait notre esprit « super-critique » (Berry, 2011), un esprit générant plus d'idées qu'il ne reçoit, mais plus pertinemment, plus d'interrogations à partir d'une multitude de lectures différentes. On devrait aussi lutter contre l'hégémonie de l'homogénéité, et en cela, ces trois « métadroits » proposés par Rouvroy et Berns, peuvent être un point de départ : « droit à l'oubli », « droit à la désobéissance », « droit de (se) rendre compte » (Rouvroy & Berns, 2010). Dans tout cela, on retrouve de nombreuses graines déjà semées par le mouvement de la psychothérapie institutionnelle, qui dans sa prise en compte aussi radicale de l'humanité du malade, peut nous aider à penser une psychiatrie à l'ère du numérique. La tâche est urgente.

Bibliographie

About BD2K | Data Science at NIH. (s. d.). Consulté 31 octobre 2017, à l'adresse <https://datascience.nih.gov/bd2k/about>

Adib, F., Mao, H., Kabelac, Z., Katabi, D., & Miller, R. C. (2015). Smart Homes that Monitor Breathing and Heart Rate (p. 837-846). ACM Press. <https://doi.org/10.1145/2702123.2702200>

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. Consulté à l'adresse <http://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596>

Anderson, C. (2008, juin 23). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. Consulté à l'adresse <https://www.wired.com/2008/06/pb-theory/>

Anderson, J. (2018, avril 11). Data engineers vs. data scientists. Consulté 1 août 2018, à l'adresse <https://www.oreilly.com/ideas/data-engineers-vs-data-scientists>

Anthes, E. (2016). Mental health: There's an app for that. *Nature News*, 532(7597), 20. <https://doi.org/10.1038/532020a>

Apple Watch Series 4 - Health. (s. d.). Consulté 14 septembre 2018, à l'adresse <https://www.apple.com/apple-watch-series-4/health/>

arXiv.org e-Print archive. (s. d.). Consulté 1 novembre 2017, à l'adresse <https://arxiv.org/>

Associated Press. (2018, janvier 30). Amazon, JPMorgan, Berkshire creating new health care company. Consulté 9 septembre 2018, à l'adresse <https://www.statnews.com/2018/01/30/amazon-jpmorgan-berkshire-health-care/>

ATLAS Collaboration. (2018). A strategy for a general search for new phenomena using data-derived signal regions and its application within the ATLAS experiment. *arXiv:1807.07447 [hep-ex]*. Consulté à l'adresse <http://arxiv.org/abs/1807.07447>

Auden, W. H. (1946). Under which lyre: a reactionary tract for the times. *Collected Poems*, 333-334.

Auffray, C., Balling, R., Barroso, I., Bencze, L., Benson, M., Bergeron, J., ... Zanetti, G. (2016). Making sense of big data in health research: Towards an EU action plan. *Genome Medicine*, 8, 71. <https://doi.org/10.1186/s13073-016-0323-y>

Bach, N. (2017, 08). Google Will Now Ask Users: « Are You Depressed? » Consulté 30 octobre 2017, à l'adresse <http://fortune.com/2017/08/24/google-depression-screening/>

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2016). Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. *JMIR Mental Health*, 3(1), e7. <https://doi.org/10.2196/mental.4984>

Battelle, J. (2005, 01). The Birth of Google. Consulté 29 octobre 2017, à l'adresse <https://www.wired.com/2005/08/battelle/>

Bedi, G., Carrillo, F., Cecchi, G. A., Slezak, D. F., Sigman, M., Mota, N. B., ... Corcoran, C. M. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia*, 1, 15030. <https://doi.org/10.1038/npjSchz.2015.30>

- Béguin, F. (2018a, janvier 26). La psychiatrie en grande souffrance. Consulté à l'adresse https://www.lemonde.fr/sciences/article/2018/01/26/le-grand-malaise-des-soignants-en-psychiatrie-on-a-le-sentiment-d-etre-maltraitant_5247248_1650684.html
- Béguin, F. (2018b, août 18). Dans les hôpitaux psychiatriques, les soignants dénoncent une situation « critique ». Consulté à l'adresse https://www.lemonde.fr/sante/article/2018/08/18/dans-les-hopitaux-psychiatriques-les-soignants-denoncent-une-situation-critique_5343762_1651302.html
- Belgard, T. G., Jankovic, I., Lowe, J. K., & Geschwind, D. H. (2014). Population structure confounds autism genetic classifier. *Molecular psychiatry*, 19(4), 405-407. <https://doi.org/10.1038/mp.2013.34>
- Belluck, P. (2017, novembre 13). First Digital Pill Approved to Worries About Biomedical 'Big Brother'. *The New York Times*. Consulté à l'adresse <https://www.nytimes.com/2017/11/13/health/digital-pill-fda.html>
- Ben-Zeev, D., Brenner, C. J., Begale, M., Duffecy, J., Mohr, D. C., & Mueser, K. T. (2014). Feasibility, Acceptability, and Preliminary Efficacy of a Smartphone Intervention for Schizophrenia. *Schizophrenia Bulletin*, 40(6), 1244-1253. <https://doi.org/10.1093/schbul/sbu033>
- Ben-Zeev, D., Brian, R., Wang, R., Wang, W., Campbell, A. T., Aung, M. S. H., ... Scherer, E. A. (2017). CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatric Rehabilitation Journal*, 40(3), 266-275. <https://doi.org/10.1037/prj0000243>
- Berry, D. M. (2011). *The philosophy of software: code and mediation in the digital age*. Basingstoke, Hampshire ; New York: Palgrave Macmillan.
- Bhugra, D., Tasman, A., Pathare, S., Priebe, S., Smith, S., Torous, J., ... Ventriglio, A. (2017). The WPA- Lancet Psychiatry Commission on the Future of Psychiatry. *The Lancet Psychiatry*, 4(10), 775-818. [https://doi.org/10.1016/S2215-0366\(17\)30333-4](https://doi.org/10.1016/S2215-0366(17)30333-4)
- Biagiante, B., Schlosser, D., Nahum, M., Woolley, J., & Vinogradov, S. (2016). Creating Live Interactions to Mitigate Barriers (CLIMB): A Mobile Intervention to Improve Social Functioning in People With Chronic Psychotic Disorders. *JMIR Mental Health*, 3(4). <https://doi.org/10.2196/mental.6671>
- Big data | CNIL. (s. d.). Consulté 23 octobre 2017, à l'adresse <https://www.cnil.fr/fr/definition/big-data>
- Boseley, S. (2016, 07). NHS to scrap single database of patients' medical details. *The Guardian*. Consulté à l'adresse <http://www.theguardian.com/technology/2016/jul/06/nhs-to-scrap-single-database-of-patients-medical-details>
- Bourdieu, P., & Wacquant, L. (2000, mai 1). La nouvelle vulgate planétaire. *Le Monde diplomatique*. Consulté à l'adresse <https://www.monde-diplomatique.fr/2000/05/BOURDIEU/2269>
- Bower, J. L., & Christensen, C. M. (1995, janvier 1). Disruptive Technologies: Catching the Wave. *Harvard Business Review*, (January–February 1995). Consulté à l'adresse <https://hbr.org/1995/01/disruptive-technologies-catching-the-wave>
- boyd, danah, & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662-679. <https://doi.org/10.1080/1369118X.2012.678878>

- Boyko, I. (s. d.). « Users. Android L (Lollipop) » by Ivan Boyko. Consulté 13 novembre 2018, à l'adresse <https://www.iconfinder.com/iconsets/users-android-l-lollipop>
- brainmap.org | Home. (s. d.). Consulté 28 mars 2018, à l'adresse <http://brainmap.org/>
- Breiman, L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199-231. <https://doi.org/10.1214/ss/1009213726>
- Brodersen, K. H., Deserno, L., Schlagenhaut, F., Lin, Z., Penny, W. D., Buhmann, J. M., & Stephan, K. E. (2014). Dissecting psychiatric spectrum disorders by generative embedding. *NeuroImage: Clinical*, 4, 98-111. <https://doi.org/10.1016/j.nicl.2013.11.002>
- Brooker, C. (2011). Black Mirror. Consulté à l'adresse <http://www.imdb.com/title/tt2085059/>
- Bucci, S., Barrowclough, C., Ainsworth, J., Machin, M., Morris, R., Berry, K., ... Haddock, G. (2018). Actissist: Proof-of-Concept Trial of a Theory-Driven Digital Intervention for Psychosis. *Schizophrenia Bulletin*, 44(5), 1070-1080. <https://doi.org/10.1093/schbul/sby032>
- Buhr, S. (2015, mai). IBM's Watson Wants To Help Pick A Therapist For You. Consulté 20 septembre 2018, à l'adresse <http://social.techcrunch.com/2015/05/10/ibms-watson-wants-to-help-pick-a-therapist-for-you/>
- Building a Safer Community With New Suicide Prevention Tools | Facebook Newsroom. (s. d.). Consulté 2 novembre 2017, à l'adresse <https://newsroom.fb.com/news/2017/03/building-a-safer-community-with-new-suicide-prevention-tools/>
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Conference on Fairness, Accountability and Transparency* (p. 77-91). Consulté à l'adresse <http://proceedings.mlr.press/v81/buolamwini18a.html>
- Cabitza, F., Rasoini, R., & Gensini, G. F. (2017). Unintended Consequences of Machine Learning in Medicine. *JAMA*, 318(6), 517-518. <https://doi.org/10.1001/jama.2017.7797>
- CAB_Solidarités. (2018, juin 28). Santé mentale et psychiatrie : Agnès Buzyn présente sa feuille de route pour changer le regard sur la santé mentale et les personnes atteintes de troubles psychiques. Consulté 21 septembre 2018, à l'adresse <https://solidarites-sante.gouv.fr/actualites/presse/communiqués-de-presse/article/sante-mentale-et-psychiatrie-agnes-buzyn-presente-sa-feuille-de-route-pour>
- Cannon, T. D., Yu, C., Addington, J., Bearden, C. E., Cadenhead, K. S., Cornblatt, B. A., ... Kattan, M. W. (2016). An Individualized Risk Calculator for Research in Prodromal Psychosis. *American Journal of Psychiatry*, 173(10), 980-988. <https://doi.org/10.1176/appi.ajp.2016.15070890>
- Carrión, R. E., Cornblatt, B. A., Burton, C. Z., Tso, I. F., Auther, A. M., Adelsheim, S., ... McFarlane, W. R. (2016). Personalized Prediction of Psychosis: External Validation of the NAPLS-2 Psychosis Risk Calculator With the EDIPPP Project. *American Journal of Psychiatry*, 173(10), 989-996. <https://doi.org/10.1176/appi.ajp.2016.15121565>
- Castro, V. M., Minnier, J., Murphy, S. N., Kohane, I., Churchill, S. E., Gainer, V., ... Smoller, J. W. (2015). Validation of Electronic Health Record Phenotyping of Bipolar Disorder and Controls. *The American journal of psychiatry*, 172(4), 363-372. <https://doi.org/10.1176/appi.ajp.2014.14030423>
- CDC. (2017, mai 5). Ten Great Public Health Achievements in the 20th Century. Consulté 11 octobre 2017, à l'adresse <https://www.cdc.gov/about/history/tengpha.htm>

CDC. (s. d.). Achievements in Public Health, 1900-1999: Changes in the Public Health System. Consulté 11 octobre 2017, à l'adresse <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm4850a1.htm>

Center for Devices and Radiological Health. (s. d.). CDRH Transparency - Evaluation of Automatic Class III Designation (De Novo) Summaries [WebContent]. Consulté 14 septembre 2018, à l'adresse <https://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHTransparency/ucm232269.htm>

Centre National de Ressources Textuelles et Lexicales. (s. d.-a). Algorithme : Etymologie de Algorithme. Consulté 11 novembre 2018, à l'adresse <http://www.cnrtl.fr/etymologie/algorithme>

Centre National de Ressources Textuelles et Lexicales. (s. d.-b). STATISTIQUE : Etymologie de STATISTIQUE. Consulté 30 octobre 2018, à l'adresse <http://www.cnrtl.fr/etymologie/statistique>

Chan, M. K., Krebs, M.-O., Cox, D., Guest, P. C., Yolken, R. H., Rahmoune, H., ... Bahn, S. (2015). Development of a blood-based molecular biomarker test for identification of schizophrenia before disease onset. *Translational Psychiatry*, 5(7), e601. <https://doi.org/10.1038/tp.2015.91>

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *New England Journal of Medicine*, 378(11), 981-983. <https://doi.org/10.1056/NEJMp1714229>

CHARGE inflammation working group, Milaneschi, Y., Lamers, F., Peyrot, W. J., Abdellaoui, A., Willemsen, G., ... Penninx, B. W. J. H. (2016). Polygenic dissection of major depression clinical heterogeneity. *Molecular Psychiatry*, 21(4), 516-522. <https://doi.org/10.1038/mp.2015.86>

Chekroud, Adam M., Gueorguieva, R., Krumholz, H. M., Trivedi, M. H., Krystal, J. H., & McCarthy, G. (2017). Reevaluating the Efficacy and Predictability of Antidepressant Treatments: A Symptom Clustering Approach. *JAMA Psychiatry*, 74(4), 370-378. <https://doi.org/10.1001/jamapsychiatry.2017.0025>

Chekroud, Adam Mourad, Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., ... Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: a machine learning approach. *The Lancet Psychiatry*, 3(3), 243-250. [https://doi.org/10.1016/S2215-0366\(15\)00471-X](https://doi.org/10.1016/S2215-0366(15)00471-X)

Chen, L. Y. (2018, juillet 3). A \$6 Billion China Startup Wants to Be the Amazon of Health Care. Consulté 12 septembre 2018, à l'adresse <https://www.bloomberg.com/news/articles/2018-07-02/a-6-billion-china-startup-wants-to-be-the-amazon-of-health-care>

Chen, Y. J., Liu, C. M., Hsu, Y. C., Lo, Y. C., Hwang, T. J., Hwu, H. G., ... Tseng, W. I. (2018). Individualized prediction of schizophrenia based on the whole-brain pattern of altered white matter tract integrity. *Human Brain Mapping*, 39(1), 575-587. <https://doi.org/10.1002/hbm.23867>

Chervenak, A., Foster, I., Kesselman, C., Salisbury, C., & Tuecke, S. (2000). The data grid: Towards an architecture for the distributed management and analysis of large scientific datasets. *Journal of Network and Computer Applications*, 23(3), 187-200. <https://doi.org/10.1006/jnca.2000.0110>

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... Greene, C. S. (2018). Opportunities And Obstacles For Deep Learning In Biology And Medicine. <https://doi.org/10.1101/142760>

- Chowkwanyun, M., Bayer, R., & Galea, S. (2018). "Precision" Public Health — Between Novelty and Hype. *New England Journal of Medicine*, 379(15), 1398-1400. <https://doi.org/10.1056/NEJMp1806634>
- Christensen, C. M., Raynor, M. E., & McDonald, R. (2015, décembre 1). What Is Disruptive Innovation? *Harvard Business Review*, (December 2015). Consulté à l'adresse <https://hbr.org/2015/12/what-is-disruptive-innovation>
- Chuang, L.-C., & Kuo, P.-H. (2017). Building a genetic risk model for bipolar disorder from genome-wide association data with random forest algorithm. *Scientific Reports*, 7, 39943. <https://doi.org/10.1038/srep39943>
- Ciarleglio, A. J., Brucato, G., Masucci, M. D., Altschuler, R., Colibazzi, T., Corcoran, C. M., ... Girgis, R. R. (2018). A predictive model for conversion to psychosis in clinical high-risk patients. *Psychological Medicine*, 1-10. <https://doi.org/10.1017/S003329171800171X>
- Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ileva, E. I., Ethridge, L. E., Pearlson, G. D., ... Tamminga, C. A. (2015). Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers. *American Journal of Psychiatry*, 173(4), 373-384. <https://doi.org/10.1176/appi.ajp.2015.14091200>
- Clinical Record Interactive Search (CRIS). (s. d.). Consulté 25 juillet 2017, à l'adresse <http://www.maudsleybrc.nihr.ac.uk/facilities/clinical-record-interactive-search-cris/>
- ClinicalTrials.gov. (s. d.). Consulté 31 octobre 2017, à l'adresse <https://clinicaltrials.gov/>
- Code de la santé publique - Article L5211-1, L5211-1 Code de la santé publique §.
- Code de la santé publique - Article R5211-1, R5211-1 Code de la santé publique §.
- Cohen, A. S., & Elvevåg, B. (2014). Automated computerized analysis of speech in psychiatric disorders. *Current Opinion in Psychiatry*, 27(3), 203-209. <https://doi.org/10.1097/YCO.0000000000000056>
- Comité consultatif national d'éthique. (2018). *Avis 129 « Contribution du Comité consultatif national d'éthique à la révision de la loi de bioéthique 2018-2019 »*. Consulté à l'adresse <http://www.ccn-ethique.fr/fr/publications/contribution-du-comite-consultatif-national-dethique-la-revision-de-la-loi-de>
- Comprendre vos droits | CNIL. (s. d.). Consulté 30 octobre 2017, à l'adresse <https://www.cnil.fr/fr/comprendre-vos-droits>
- Conférence Mondiale sur les Politiques Culturelles (MONDIACULT): rapport final. (1982). UNESCO. Consulté à l'adresse http://www.unesco.org/ulis/cgi-bin/ulis.pl?catno=52505&set=005BD99586_3_303&gp=1&lin=1&ll=1
- Copeland, B. J. (2017). The Modern History of Computing. In E. N. Zalta (Éd.), *The Stanford Encyclopedia of Philosophy* (Winter 2017). Metaphysics Research Lab, Stanford University. Consulté à l'adresse <https://plato.stanford.edu/archives/win2017/entries/computing-history/>
- Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural Language Processing of Social Media as Screening for Suicide Risk. *Biomedical Informatics Insights*, 10, 1178222618792860. <https://doi.org/10.1177/1178222618792860>
- Coupechoux, P., & Delion, P. (2014). *Un homme comme vous: essai sur l'humanité de la folie*. Paris: Éditions du Seuil.

- Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, *11*, 126. <https://doi.org/10.1186/1741-7015-11-126>
- Dagum, P. (2018). Digital biomarkers of cognitive function. *Npj Digital Medicine*, *1*(1). <https://doi.org/10.1038/s41746-018-0018-4>
- Data Harm Record. (2017, décembre 6). Consulté 4 novembre 2018, à l'adresse <https://datajusticelab.org/data-harm-record/>
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *ICWSM*, *13*, 1-10.
- De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016). Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media (p. 2098-2110). ACM Press. <https://doi.org/10.1145/2858036.2858207>
- Débat public sur les algorithmes et l'intelligence artificielle : programme des évènements à venir | CNIL. (s. d.). Consulté 10 octobre 2017, à l'adresse <https://www.cnil.fr/fr/debat-public-sur-les-algorithmes-et-lintelligence-artificielle-programme-des-evenements-venir>
- DeBattista, C., Kinrys, G., Hoffman, D., Goldstein, C., Zajacka, J., Kocsis, J., ... Fava, M. (2011). The use of referenced-EEG (rEEG) in assisting medication selection for the treatment of depression. *Journal of Psychiatric Research*, *45*(1), 64-75. <https://doi.org/10.1016/j.jpsychires.2010.05.009>
- Deleuze, G. (2007). « Post-scriptum sur les sociétés de contrôle ». In *Pourparlers, 1972-1990*. Paris: Editions de Minuit.
- Delion, P. (2014). La psychothérapie institutionnelle : d'où vient-elle et où va-t-elle ? *Empan*, (96), 104-112. <https://doi.org/10.3917/empa.096.0104>
- Deutsch, D. (2012, octobre). How close are we to creating artificial intelligence? Consulté 21 octobre 2018, à l'adresse <https://aeon.co/essays/how-close-are-we-to-creating-artificial-intelligence>
- Dobbs, D. (2017, août). The Smartphone Psychiatrist. *The Atlantic*. Consulté à l'adresse <https://www.theatlantic.com/magazine/archive/2017/07/the-smartphone-psychiatrist/528726/>
- Donker, T., Blankers, M., Hedman, E., Ljótsson, B., Petrie, K., & Christensen, H. (2015). Economic evaluations of Internet interventions for mental health: a systematic review. *Psychological Medicine*, *45*(16), 3357-3376. <https://doi.org/10.1017/S0033291715001427>
- Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., ... Schlaggar, B. L. (2010). Prediction of Individual Brain Maturity Using fMRI. *Science*, *329*(5997), 1358-1361. <https://doi.org/10.1126/science.1194144>
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., ... Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23*(1), 28-38. <https://doi.org/10.1038/nm.4246>
- Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, *98*(24), 13763-13768. <https://doi.org/10.1073/pnas.231499798>
- Editorial. (2018, août 18). La psychiatrie, un secteur en état d'urgence. Consulté à l'adresse https://www.lemonde.fr/idees/article/2018/08/18/la-psychiatrie-un-secteur-en-etat-d-urgence_5343765_3232.html

Ellevåg, B., Cohen, A. S., Wolters, M. K., Whalley, H. C., Gountouna, V.-E., Kuznetsova, K. A., ... Nicodemus, K. K. (2016). An examination of the language construct in NIMH's research domain criteria: Time for reconceptualization! *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, n/a-n/a. <https://doi.org/10.1002/ajmg.b.32438>

Employers - Spring Health. (s. d.). Consulté 26 octobre 2018, à l'adresse <https://www.springhealth.com/employers/>

Ensmenger, N. (2012). The Digital Construction of Technology: Rethinking the History of Computers in Society. *Technology and Culture*, 53(4), 753-776. <https://doi.org/10.1353/tech.2012.0126>

Esmail, S. (2015). Mr. Robot. Consulté à l'adresse <http://www.imdb.com/title/tt4158110/>

Espérance de vie - Mortalité—Tableaux de l'Économie Française | Insee. (s. d.). Consulté 15 octobre 2017, à l'adresse <https://www.insee.fr/fr/statistiques/1906668?sommaire=1906743>

Etkin, A., Patenaude, B., Song, Y. J. C., Usherwood, T., Rekshan, W., Schatzberg, A. F., ... Williams, L. M. (2015). A Cognitive–Emotional Biomarker for Predicting Remission with Antidepressant Medications: A Report from the iSPOT-D Trial. *Neuropsychopharmacology*, 40(6), 1332-1342. <https://doi.org/10.1038/npp.2014.333>

EU Clinical Trials Register. (s. d.). Consulté 12 février 2018, à l'adresse <https://www.clinicaltrialsregister.eu/about.html>

Executive Office of the President. (2014). *Big Data: Seizing Opportunities, Preserving Values*.

Executive Office of the President, & President's Council of Advisors on Science and Technology. (2014). *Big Data and Privacy: A Technological Perspective*. Executive Office of the President. Consulté à l'adresse https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf

Faghmous, J. H., & Kumar, V. (2014). A Big Data Guide to Understanding Climate Change: The Case for Theory-Guided Data Science. *Big Data*, 2(3), 155-163. <https://doi.org/10.1089/big.2014.0026>

Farb, N. A. S., Anderson, A. K., Bloch, R. T., & Segal, Z. V. (2011). Mood-Linked Responses in Medial Prefrontal Cortex Predict Relapse in Patients with Recurrent Unipolar Depression. *Biological Psychiatry*, 70(4), 366-372. <https://doi.org/10.1016/j.biopsych.2011.03.009>

Fauw, J. D., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., ... Ronneberger, O. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 24(9), 1342-1350. <https://doi.org/10.1038/s41591-018-0107-6>

Firth, J., Torous, J., Nicholas, J., Carney, R., Prapat, A., Rosenbaum, S., & Sarris, J. (2017). The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*, 16(3), 287-298. <https://doi.org/10.1002/wps.20472>

Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S., & Sarris, J. (2017). Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 218, 15-22. <https://doi.org/10.1016/j.jad.2017.04.046>

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent

(Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, 4(2), e19.
<https://doi.org/10.2196/mental.7785>

Fiveash, K. (2017, avril 18). NHS finally launches its Digital Apps Library—only one app “approved” so far. Consulté 26 octobre 2018, à l'adresse <https://arstechnica.com/tech-policy/2017/04/nhs-digital-apps-library/>

Form S-1 - Registration Statement under the Securities Act of 1933 - Facebook, Inc. (2012, février 1). Consulté à l'adresse <https://www.sec.gov/Archives/edgar/data/1326801/000119312512034517/d287954ds1.htm>

Fortune 500 Companies 2017: Who Made the List. (s. d.). Consulté 22 octobre 2017, à l'adresse <http://fortune.com/fortune500/list/>

Foulquié, P., Voillot, P., Faviez, C., Mebarki, A., Chen, X., Texier, N., & Schück, S. (2018). Étude de l'usage du méthylphénidate sur les réseaux sociaux. *Revue d'Épidémiologie et de Santé Publique*, 66, S225. <https://doi.org/10.1016/j.respe.2018.04.030>

Frances, A. (2013). The New Crisis in Confidence in Psychiatric Diagnosis. *Annals of Internal Medicine*. <https://doi.org/10.7326/0003-4819-159-3-201308060-00655>

Freepik. (s. d.). *Analytics free vector icon designed by Freepik*. Consulté à l'adresse https://www.flaticon.com/free-icon/analytics_413708#term=data%20analytics&page=1&position=25

Freud, S., Strachey, J., Cixous, H., & Dennomé, R. (1976). Fiction and Its Phantoms: A Reading of Freud's Das Unheimliche (The « Uncanny »). *New Literary History*, 7(3), 525-645.
<https://doi.org/10.2307/468561>

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4), 1273-1302. [https://doi.org/10.1016/S1053-8119\(03\)00202-7](https://doi.org/10.1016/S1053-8119(03)00202-7)

Ganascia, J.-G. (2010). Epistemology of AI Revisited in the Light of the Philosophy of Information. *Knowledge, Technology & Policy*, 23(1), 57-73. <https://doi.org/10.1007/s12130-010-9101-0>

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumeé III, H., & Crawford, K. (2018). Datasheets for Datasets. *arXiv:1803.09010 [cs]*. Consulté à l'adresse <http://arxiv.org/abs/1803.09010>

Gibney, E. (2017). Self-taught AI is best yet at strategy game Go. *Nature News*.
<https://doi.org/10.1038/nature.2017.22858>

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, 5.
<https://doi.org/10.7554/eLife.11305>

Gillan, C. M., & Whelan, R. (2017). What big data can do for treatment in psychiatry. *Current Opinion in Behavioral Sciences*, 18(Supplement C), 34-42.
<https://doi.org/10.1016/j.cobeha.2017.07.003>

Ginger.io. (s. d.). Consulté 2 novembre 2017, à l'adresse <https://ginger.io/>

Glatt, S. J., Tsuang, M. T., Winn, M., Chandler, S. D., Collins, M., Lopez, L., ... Courchesne, E. (2012). Blood-Based Gene Expression Signatures of Autistic Infants and Toddlers. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(9), 934-44.e2.
<https://doi.org/10.1016/j.jaac.2012.07.007>

Glossaire | CNIL. (s. d.). Consulté 6 novembre 2017, à l'adresse <https://www.cnil.fr/fr/glossaire>

Google Trends. (s. d.). Consulté 23 octobre 2017, à l'adresse [/trends/explore](https://trends.google.com/trends/explore)

Grenouilloux, A. (2017). Psychiatrie phénoménologique, médecine de la personne et big data. *Annales Médico-psychologiques, revue psychiatrique*, 175(10), 882-888. <https://doi.org/10.1016/j.amp.2017.02.014>

Grisanzio, K. A., Goldstein-Piekarski, A. N., Wang, M. Y., Rashed Ahmed, A. P., Samara, Z., & Williams, L. M. (2018). Transdiagnostic Symptom Clusters and Associations With Brain, Behavior, and Daily Function in Mood, Anxiety, and Trauma Disorders. *JAMA Psychiatry*, 75(2), 201. <https://doi.org/10.1001/jamapsychiatry.2017.3951>

Grist, R., Porter, J., & Stallard, P. (2017). Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. *Journal of Medical Internet Research*, 19(5), e176. <https://doi.org/10.2196/jmir.7332>

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... Webster, D. R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 316(22), 2402-2410. <https://doi.org/10.1001/jama.2016.17216>

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18, 43-49. <https://doi.org/10.1016/j.cobeha.2017.07.005>

Gustafson, D. H., McTavish, F. M., Chih, M.-Y., Atwood, A. K., Johnson, R. A., Boyle, M. G., ... Shah, D. (2014). A Smartphone Application to Support Recovery From Alcoholism: A Randomized Clinical Trial. *JAMA Psychiatry*, 71(5), 566-572. <https://doi.org/10.1001/jamapsychiatry.2013.4642>

Haenisch, F., Cooper, J. D., Reif, A., Kittel-Schneider, S., Steiner, J., Leweke, F. M., ... Bahn, S. (2016). Towards a blood-based diagnostic panel for bipolar disorder. *Brain, Behavior, and Immunity*, 52, 49-57. <https://doi.org/10.1016/j.bbi.2015.10.001>

Hagstrom, S. (2014, septembre 3). The FAIR Data Principles. Consulté 12 février 2018, à l'adresse <https://www.force11.org/group/fairgroup/fairprinciples>

Haigh, T. (2011). The history of information technology. *Annual Review of Information Science and Technology*, 45(1), 431-487. <https://doi.org/10.1002/aris.2011.1440450116>

Hastie, T., Tibshirani, R., & Friedman, J. H. (2017). *The elements of statistical learning: data mining, inference, and prediction* (Second edition, corrected at 12th printing 2017). New York: Springer.

Haute Autorité de santé. (2016). *Référentiel de bonnes pratiques sur les applications et les objets connectés en santé (Mobile Health ou mHealth)* (Évaluation et amélioration des pratiques). Haute Autorité de santé. Consulté à l'adresse https://www.has-sante.fr/portail/upload/docs/application/pdf/2016-11/has_ref_apps_oc.pdf

Hayden, E. C. (2013). Privacy protections: The genome hacker. *Nature News*, 497(7448), 172. <https://doi.org/10.1038/497172a>

Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., ... Gu, C. H. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, 542(7641), 348-351. <https://doi.org/10.1038/nature21369>

Health Pilot » Big Data Europe. (s. d.). Consulté 1 novembre 2017, à l'adresse <https://www.big-data-europe.eu/pilot-health/>

Heinzle, J., & Stephan, K. E. (2018). Chapter 5 - Dynamic Causal Modeling and Its Application to Psychiatric Disorders. In A. Anticevic & J. D. Murray (Éd.), *Computational Psychiatry* (p. 117-144). Academic Press. <https://doi.org/10.1016/B978-0-12-809825-7.00005-5>

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., ... Zwitter, A. (2017, février 25). Will Democracy Survive Big Data and Artificial Intelligence? Consulté 19 octobre 2017, à l'adresse <https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/>

Hilbert, M., & Lopez, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60-65. <https://doi.org/10.1126/science.1200970>

Hilbert, Martin. (2016). Big Data for Development: A Review of Promises and Challenges. *Development Policy Review*, 34(1), 135-174. <https://doi.org/10.1111/dpr.12142>

Hirsch, I. B., Martinez, J., Dorsey, E. R., Finken, G., Fleming, A., Gropp, C., ... Papapetropoulos, S. (2017). Incorporating Site-less Clinical Trials Into Drug Development: A Framework for Action. *Clinical Therapeutics*, 39(5), 1064-1076. <https://doi.org/10.1016/j.clinthera.2017.03.018>

History of the Framingham Heart Study. (s. d.). Consulté 19 octobre 2017, à l'adresse <http://www.framinghamheartstudy.org/about-fhs/history.php>

Hollis, C., Falconer, C. J., Martin, J. L., Whittington, C., Stockton, S., Glazebrook, C., & Davies, E. B. (2017). Annual Research Review: Digital health interventions for children and young people with mental health problems - a systematic and meta-review. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 58(4), 474-503. <https://doi.org/10.1111/jcpp.12663>

Home - NHS Apps Library. (s. d.). Consulté 26 octobre 2018, à l'adresse https://apps.beta.nhs.uk/category/mental_health/

Iatropoulou, K. (s. d.). Open Access Policies and Mandates | Open Access. Consulté 12 février 2018, à l'adresse <https://www.openaire.eu/oa-policies-mandates>

IBM News Room. (2015, mai 5). An Ecosystem of Innovators [CTB10]. Consulté 20 septembre 2018, à l'adresse <https://www-03.ibm.com/press/us/en/pressrelease/46753.wss>

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *American Journal of Psychiatry*, 167(7), 748-751. <https://doi.org/10.1176/appi.ajp.2010.09091379>

International Consortium on Lithium Genetics (ConLi+Gen), Amare, A. T., Schubert, K. O., Hou, L., Clark, S. R., Papiol, S., ... Baune, B. T. (2018). Association of Polygenic Score for Schizophrenia and HLA Antigen and Inflammation Genes With Response to Lithium in Bipolar Affective Disorder: A Genome-Wide Association Study. *JAMA Psychiatry*, 75(1), 65-74. <https://doi.org/10.1001/jamapsychiatry.2017.3433>

InVS/Santé publique France. (s. d.). Bases de données et outils / Publications et outils / Accueil. Consulté 8 novembre 2017, à l'adresse <http://invs.santepubliquefrance.fr/Publications-et-outils/Bases-de-donnees-et-outils>

Jackson, R. G., Patel, R., Jayatilleke, N., Kolliakou, A., Ball, M., Gorrell, G., ... Stewart, R. (2017). Natural language processing to extract symptoms of severe mental illness from clinical text: the

Clinical Record Interactive Search Comprehensive Data Extraction (CRIS-CODE) project. *BMJ Open*, 7(1), e012012. <https://doi.org/10.1136/bmjopen-2016-012012>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 103). New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4614-7138-7>

Jonas, E., & Kording, K. P. (2017). Could a Neuroscientist Understand a Microprocessor? *PLOS Computational Biology*, 13(1), e1005268. <https://doi.org/10.1371/journal.pcbi.1005268>

Just, M. A., Cherkassky, V. L., Buchweitz, A., Keller, T. A., & Mitchell, T. M. (2014). Identifying Autism from Neural Representations of Social Interactions: Neurocognitive Markers of Autism. *PLOS ONE*, 9(12), e113879. <https://doi.org/10.1371/journal.pone.0113879>

Just, M. A., Pan, L., Cherkassky, V. L., McMakin, D. L., Cha, C., Nock, M. K., & Brent, D. (2017). Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. *Nature Human Behaviour*, 1. <https://doi.org/10.1038/s41562-017-0234-y>

Kaggle: Your Home for Data Science. (s. d.). Consulté 1 novembre 2017, à l'adresse <https://www.kaggle.com/>

Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17(12), 1174-1179. <https://doi.org/10.1038/mp.2012.105>

Kawasaki, Y., Suzuki, M., Kherif, F., Takahashi, T., Zhou, S.-Y., Nakamura, K., ... Kurachi, M. (2007). Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *NeuroImage*, 34(1), 235-242. <https://doi.org/10.1016/j.neuroimage.2006.08.018>

Kendell, R., & Jablensky, A. (2003). Distinguishing Between the Validity and Utility of Psychiatric Diagnoses. *American Journal of Psychiatry*, 160(1), 4-12. <https://doi.org/10.1176/appi.ajp.160.1.4>

Kessler, R. C., Warner, C. H., Ivany, C., Petukhova, M. V., Rose, S., Bromet, E. J., ... Ursano, R. J. (2015). Predicting Suicides After Psychiatric Hospitalization in US Army Soldiers: The Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *JAMA Psychiatry*, 72(1), 49-57. <https://doi.org/10.1001/jamapsychiatry.2014.1754>

Key Changes with the General Data Protection Regulation. (s. d.). Consulté 14 septembre 2017, à l'adresse <http://eugdpr.org/the-regulation.html>

Khare, R., Ruth, B. J., Miller, M., Tucker, J., Utidjian, L. H., Razzaghi, H., ... Bailey, L. C. (2018). Predicting Causes of Data Quality Issues in a Clinical Data Research Network. *AMIA Joint Summits on Translational Science Proceedings. AMIA Joint Summits on Translational Science, 2017*, 113-121.

Khare, R., Utidjian, L., Ruth, B. J., Kahn, M. G., Burrows, E., Marsolo, K., ... Bailey, L. C. (2017). A longitudinal analysis of data quality in a large pediatric data research network. *Journal of the American Medical Informatics Association: JAMIA*, 24(6), 1072-1079. <https://doi.org/10.1093/jamia/ocx033>

Khoury, M. J., & Ioannidis, J. P. A. (2014). Big data meets public health. *Science*, 346(6213), 1054-1055. <https://doi.org/10.1126/science.aaa2709>

Kim, Y.-J., Jang, H. M., Lee, Y., Lee, D., & Kim, D.-J. (2018). Effects of Internet and Smartphone Addictions on Depression and Anxiety Based on Propensity Score Matching Analysis. *International*

Journal of Environmental Research and Public Health, 15(5).
<https://doi.org/10.3390/ijerph15050859>

Koomey, J., Berard, S., Sanchez, M., & Wong, H. (2011). Implications of Historical Trends in the Electrical Efficiency of Computing. *IEEE Annals of the History of Computing*, 33(3), 46-54.
<https://doi.org/10.1109/MAHC.2010.28>

Koutsouleris, N., Borgwardt, S., Meisenzahl, E. M., Bottlender, R., Möller, H.-J., & Riecher-Rössler, A. (2012). Disease Prediction in the At-Risk Mental State for Psychosis Using Neuroanatomical Biomarkers: Results From the FePsy Study. *Schizophrenia Bulletin*, 38(6), 1234-1246.
<https://doi.org/10.1093/schbul/sbr145>

Koutsouleris, N., Davatzikos, C., Bottlender, R., Patschurek-Kliche, K., Scheuerecker, J., Decker, P., ... Meisenzahl, E. M. (2012). Early Recognition and Disease Prediction in the At-Risk Mental States for Psychosis Using Neurocognitive Pattern Classification. *Schizophrenia Bulletin*, 38(6), 1200-1215. <https://doi.org/10.1093/schbul/sbr037>

Koutsouleris, N., Kahn, R. S., Chekroud, A. M., Leucht, S., Falkai, P., Wobrock, T., ... Hasan, A. (2016). Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *The Lancet Psychiatry*, 3(10), 935-946.
[https://doi.org/10.1016/S2215-0366\(16\)30171-7](https://doi.org/10.1016/S2215-0366(16)30171-7)

Koutsouleris, N., Riecher-Rössler, A., Meisenzahl, E. M., Smieskova, R., Studerus, E., Kambeitz-Illankovic, L., ... Borgwardt, S. (2015). Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. *Schizophrenia Bulletin*, 41(2), 471-482.
<https://doi.org/10.1093/schbul/sbu078>

Kupfer, D. J., & Regier, D. A. (2011). Neuroscience, clinical evidence, and the future of psychiatric classification in DSM-5. *The American Journal of Psychiatry*, 168(7), 672-674.
<https://doi.org/10.1176/appi.ajp.2011.11020219>

Kuss, D. J., Harkin, L., Kanjo, E., & Billieux, J. (2018). Problematic Smartphone Use: Investigating Contemporary Experiences Using a Convergent Design. *International Journal of Environmental Research and Public Health*, 15(1). <https://doi.org/10.3390/ijerph15010142>

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
<https://doi.org/10.1038/nature14539>

Lepore, J. (2014, juin 16). What the Gospel of Innovation Gets Wrong. Consulté à l'adresse <https://www.newyorker.com/magazine/2014/06/23/the-disruption-machine>

Les missions | CNIL. (s. d.). Consulté 12 octobre 2017, à l'adresse <https://www.cnil.fr/fr/les-missions>

Les principes clés de la protection des données personnelles | CNIL. (s. d.). Consulté 7 novembre 2017, à l'adresse <https://www.cnil.fr/fr/comprendre-vos-obligations/les-principes-cles>

Librenza-Garcia, D., Kotzian, B. J., Yang, J., Mwangi, B., Cao, B., Pereira Lima, L. N., ... Passos, I. C. (2017). The impact of machine learning techniques in the study of bipolar disorder: A systematic review. *Neuroscience & Biobehavioral Reviews*, 80(Supplement C), 538-554.
<https://doi.org/10.1016/j.neubiorev.2017.07.004>

Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., ... Stumpe, M. C. (2017). Detecting Cancer Metastases on Gigapixel Pathology Images. *arXiv:1703.02442 [cs]*. Consulté à l'adresse <http://arxiv.org/abs/1703.02442>

Live better, together! | PatientsLikeMe. (s. d.). Consulté 2 novembre 2017, à l'adresse <https://www.patientslikeme.com/>

Lohr, S. (2013, janvier 2). The Origins of « Big Data »: An Etymological Detective Story - The New York Times. Consulté 11 octobre 2017, à l'adresse <https://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/>

Lohr, S. (2014, août 17). For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights. *The New York Times*. Consulté à l'adresse <https://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html>

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

LOI n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé - Article 193, 2016-41 § (2016).

Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical Informatics Insights*, 8, 1-10. <https://doi.org/10.4137/BII.S31559>

Lupton, D. (2014a). Critical Perspectives on Digital Health Technologies. *Sociology Compass*, 8(12), 1344-1359. <https://doi.org/10.1111/soc4.12226>

Lupton, D. (2014b). The commodification of patient opinion: the digital patient experience economy in the age of big data. *Sociology of Health & Illness*, 36(6), 856-869. <https://doi.org/10.1111/1467-9566.12109>

Lupton, D. (2016). *The Quantified Self* (1 edition). Cambridge, UK: Polity.

Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 1(2), 2053951714541861. <https://doi.org/10.1177/2053951714541861>

Mahoney, M. S. (2005). The histories of computing(s). *Interdisciplinary Science Reviews*, 30(2), 119-135. <https://doi.org/10.1179/030801805X25927>

Maier, R., Moser, G., Chen, G.-B., Ripke, S., Coryell, W., Potash, J. B., ... Zöllner, S. (2015). Joint Analysis of Psychiatric Disorders Increases Accuracy of Risk Prediction for Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. *The American Journal of Human Genetics*, 96(2), 283-294. <https://doi.org/10.1016/j.ajhg.2014.12.006>

Mäkinen, V.-P. (2018). Challenges in conducting genetic analyses based on data-driven classification of major depressive disorder. *Molecular Psychiatry*, 23(3), 494-494. <https://doi.org/10.1038/mp.2016.205>

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity* | McKinsey & Company. Consulté à l'adresse <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>

Markoff, J. (2007, janvier 10). Apple Introduces Innovative Cellphone. *The New York Times*. Consulté à l'adresse <https://www.nytimes.com/2007/01/10/technology/10apple.html>

Marom, S. (2009). On the Precarious Path of Reverse Neuro-Engineering. *Frontiers in Computational Neuroscience*, 3. <https://doi.org/10.3389/neuro.10.005.2009>

Marsh, S. (2017, mai 14). NHS cancer patients hit by treatment delays after cyber-attack. *The Guardian*. Consulté à l'adresse <http://www.theguardian.com/society/2017/may/14/nhs-cancer-patients-treatment-delays-cyber-attack>

Martínez-Pérez, B., de la Torre-Díez, I., & López-Coronado, M. (2013). Mobile Health Applications for the Most Prevalent Conditions by the World Health Organization: Review and Analysis. *Journal of Medical Internet Research*, 15(6). <https://doi.org/10.2196/jmir.2600>

Martos, S. (2017, 09). Intelligence artificielle : une stratégie nationale pour janvier, la santé au menu. Consulté 25 octobre 2017, à l'adresse https://www.lequotidiendumedecin.fr/actualites/article/2017/09/08/intelligence-artificielle-une-strategie-nationale-pour-janvier-la-sante-au-menu_850092

McCoy, T. H., Castro, V. M., Cagan, A., Roberson, A. M., Kohane, I. S., & Perlis, R. H. (2015). Sentiment Measured in Hospital Discharge Notes Is Associated with Readmission and Mortality Risk: An Electronic Health Record Study. *PLOS ONE*, 10(8), e0136341. <https://doi.org/10.1371/journal.pone.0136341>

McCoy, T. H., Castro, V. M., Roberson, A. M., Snapper, L. A., & Perlis, R. H. (2016). Improving Prediction of Suicide and Accidental Death After Discharge From General Hospitals With Natural Language Processing. *JAMA Psychiatry*, 73(10), 1064-1071. <https://doi.org/10.1001/jamapsychiatry.2016.2172>

McIntosh, A. M., Stewart, R., John, A., Smith, D. J., Davis, K., Sudlow, C., ... Porteous, D. J. (2016). Data science for mental health: a UK perspective on a global challenge. *The Lancet Psychiatry*, 3(10), 993-998. [https://doi.org/10.1016/S2215-0366\(16\)30089-X](https://doi.org/10.1016/S2215-0366(16)30089-X)

Medibio – Innovation In Mind. (s. d.). Consulté 9 octobre 2017, à l'adresse <http://medibio.com.au/>

Mental Health Research Charity. (s. d.). Consulté 2 novembre 2017, à l'adresse <https://www.mqmentalhealth.org/>

Meystre, S. M., Savova, G. K., Kipper-Schuler, K. C., & Hurdle, J. F. (2008). Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. *IMIA Yearbook*, 128-144.

Micoulaud-Franchi, J.-A., Quiles, C., Batail, J.-M., Lancon, C., Masson, M., Dumas, G., & Cermolacce, M. (2018). Making psychiatric semiology great again: A semiologic, not nosologic challenge. *L'Encéphale*, 44(4), 343-353. <https://doi.org/10.1016/j.encep.2018.01.007>

Micoulaud-Franchi, Jean-Arthur, Dumas, G., Quiles, C., & Vion-Dury, J. (2017). De la clinique au « terrain fétide et palpitant de la vie » : une mise en perspective psychiatrique de la physiologie clinique. *Annales Médico-psychologiques, revue psychiatrique*, 175(1), 70-85. <https://doi.org/10.1016/j.amp.2016.11.004>

Ministère des Solidarités et de la Santé. (2018a, juin 12). Agnès Buzyn lance la mission de préfiguration du « Health Data Hub » un laboratoire d'exploitation des données de santé. Consulté 26 juillet 2018, à l'adresse <http://solidarites-sante.gouv.fr/actualites/presse/communiqués-de-presse/article/agnes-buzyn-lance-la-mission-de-prefiguration-du-health-data-hub-un-laboratoire>

Ministère des Solidarités et de la Santé. (2018b, septembre 18). Ma santé 2022 : un engagement collectif. Consulté 2 novembre 2018, à l'adresse <https://solidarites-sante.gouv.fr/actualites/presse/dossiers-de-presse/article/ma-sante-2022-un-engagement-collectif>

Mission Villani sur l'IA. (s. d.). AI for humanity. Consulté 17 septembre 2018, à l'adresse <https://www.aiforhumanity.fr>

Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting Human Brain Activity Associated with the Meanings of Nouns. *Science*, 320(5880), 1191-1195. <https://doi.org/10.1126/science.1152876>

Mota, Natália B., Furtado, R., Maia, P. P. C., Copelli, M., & Ribeiro, S. (2014). Graph analysis of dream reports is especially informative about psychosis. *Scientific Reports*, 4. <https://doi.org/10.1038/srep03691>

Mota, Natalia B., Vasconcelos, N. A. P., Lemos, N., Pieretti, A. C., Kinouchi, O., Cecchi, G. A., ... Ribeiro, S. (2012). Speech Graphs Provide a Quantitative Measure of Thought Disorder in Psychosis. *PLoS ONE*, 7(4), e34928. <https://doi.org/10.1371/journal.pone.0034928>

Mueller, N. E., Panch, T., Macias, C., Cohen, B. M., Ongur, D., & Baker, J. T. (2018). Using Smartphone Apps to Promote Psychiatric Rehabilitation in a Peer-Led Community Support Program: Pilot Study. *JMIR Mental Health*, 5(3), e10092. <https://doi.org/10.2196/10092>

Murdoch, T. B., & Detsky, A. S. (2013). The Inevitable Application of Big Data to Health Care. *JAMA*, 309(13), 1351-1352. <https://doi.org/10.1001/jama.2013.393>

Mwangi, B., Wu, M.-J., Cao, B., Passos, I. C., Lavagnino, L., Keser, Z., ... Soares, J. C. (2016). Individualized Prediction and Clinical Staging of Bipolar Disorders Using Neuroanatomical Biomarkers. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1(2), 186-194. <https://doi.org/10.1016/j.bpsc.2016.01.001>

National Institutes of Health (NIH) - All Of Us | National Institutes of Health (NIH). (s. d.). Consulté 9 octobre 2017, à l'adresse <https://allofus.nih.gov/>

National Research Council (U.S.) (Éd.). (2011). *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*. Washington, D.C: National Academies Press.

Naughton, J. (2017, juillet 9). Giving Google our private NHS data is simply illegal | John Naughton. *The Guardian*. Consulté à l'adresse <http://www.theguardian.com/commentisfree/2017/jul/09/giving-google-private-nhs-data-is-simply-illegal>

Nicholas, J., Fogarty, A. S., Boydell, K., & Christensen, H. (2017). The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder. *Journal of Medical Internet Research*, 19(4), e105. <https://doi.org/10.2196/jmir.7273>

Niculescu, A. B., Levey, D. F., Phalen, P. L., Le-Niculescu, H., Dainton, H. D., Jain, N., ... Salomon, D. R. (2015). Understanding and predicting suicidality using a combined genomic and clinical risk assessment approach. *Molecular Psychiatry*, 20(11), 1266-1285. <https://doi.org/10.1038/mp.2015.112>

Nieman, D. H., Ruhrmann, S., Dragt, S., Soen, F., van Tricht, M. J., Koelman, J. H. T. . M., ... de Haan, L. (2014). Psychosis Prediction: Stratification of Risk Estimation With Information-Processing and Premorbid Functioning Variables. *Schizophrenia Bulletin*, 40(6), 1482-1490. <https://doi.org/10.1093/schbul/sbt145>

Nieuwenhuis, M., van Haren, N. E. M., Hulshoff Pol, H. E., Cahn, W., Kahn, R. S., & Schnack, H. G. (2012). Classification of schizophrenia patients and healthy controls from structural MRI scans

in two large independent samples. *NeuroImage*, 61(3), 606-612.
<https://doi.org/10.1016/j.neuroimage.2012.03.079>

NIH Data Sharing Information. (s. d.). Consulté 12 février 2018, à l'adresse
https://grants.nih.gov/grants/policy/data_sharing/

NIH Fact Sheets - Human Genome Project. (s. d.). Consulté 15 octobre 2017, à l'adresse
<https://report.nih.gov/NIHfactsheets/ViewFactSheet.aspx?csid=45>

NIH VideoCast - Artificial Intelligence: Collecting Data to Maximize Potential. (2018). Consulté à l'adresse <https://videocast.nih.gov/summary.asp?live=28053&bhcp=1>

NIMH » RDoC Matrix. (s. d.). Consulté 30 octobre 2017, à l'adresse
<https://www.nimh.nih.gov/research-priorities/rdoc/constructs/rdoc-matrix.shtml>

Olfson, M., Marcus, S. C., & Bridge, J. A. (2014). Focusing Suicide Prevention on Periods of High Risk. *JAMA*, 311(11), 1107-1108. <https://doi.org/10.1001/jama.2014.501>

O'Neil, C. (2016a). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (1 édition). New York: Crown.

O'Neil, C. (2016b, septembre 1). How algorithms rule our working lives | Cathy O'Neil. *The Guardian*. Consulté à l'adresse <https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives>

Online Therapy & Free Counseling, Someone To Talk To | 7 Cups. (s. d.). Consulté 27 octobre 2018, à l'adresse <https://www.7cups.com/>

Onnela, J.-P., & Rauch, S. L. (2016). Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. *Neuropsychopharmacology*, 41(7), 1691-1696.
<https://doi.org/10.1038/npp.2016.7>

Open PHACTS. (s. d.). Consulté 1 novembre 2017, à l'adresse <http://www.openphacts.org/>

OpenAIRE. (s. d.). Consulté 31 octobre 2017, à l'adresse <https://www.openaire.eu/>

Oquendo, M. A., Baca-Garcia, E., Artés-Rodríguez, A., Perez-Cruz, F., Galfalvy, H. C., Blasco-Fontecilla, H., ... Duan, N. (2012). Machine learning and data mining: strategies for hypothesis generation. *Molecular Psychiatry*, 17(10), 956-959. <https://doi.org/10.1038/mp.2011.173>

Organisation mondiale de la santé. (2009). *Classification statistique internationale des maladies et des problèmes de santé connexes CIM-10. 1, 1, .* Genève: Organisation mondiale de la santé.

Oury, J. (2005). *Le collectif: le séminaire de Sainte-Anne*. Nîmes: Champ social.

Oury, J. (2006a). Reims. In *Aux limites du sujet* (p. 135-139). ERES. Consulté à l'adresse <https://www.cairn.info/aux-limites-du-sujet--9782749205892-page-135.htm>

Oury, J. (2006b). Structure et vie quotidienne. In *Psychose, vie quotidienne* (p. 21-26). ERES. Consulté à l'adresse <https://www.cairn.info/psychose-vie-quotidienne-et-psychotherapie-institu--9782865865956-page-21.htm>

Oury, J. (2007a). Le corps et ses entours : la fonction scribe. In *Corps, psychose et institution* (p. 21-27). ERES. Consulté à l'adresse <https://www.cairn.info/corps-psychose-et-institution--2749200482-page-21.htm>

- Oury, J. (2007b). Psychanalyse, psychiatrie et psychothérapie institutionnelles. *VST - Vie sociale et traitements*, (95), 110-125. <https://doi.org/10.3917/vst.095.0110>
- OWKIN - Data Science and Artificial Intelligence for Medical Predictions. (s. d.). Consulté 26 juillet 2018, à l'adresse <https://owkin.com>
- Pankow, G. (2009). *L'homme et sa psychose*. Paris: Editions Flammarion.
- Passos, I. C., Mwangi, B., & Kapczinski, F. (2016). Big data analytics and machine learning: 2015 and beyond. *The Lancet Psychiatry*, 3(1), 13-15. [https://doi.org/10.1016/S2215-0366\(15\)00549-0](https://doi.org/10.1016/S2215-0366(15)00549-0)
- Patel, M. J., Khalaf, A., & Aizenstein, H. J. (2016). Studying depression using imaging and machine learning methods. *NeuroImage: Clinical*, 10(Supplement C), 115-123. <https://doi.org/10.1016/j.nicl.2015.11.003>
- Perera, G., Broadbent, M., Callard, F., Chang, C.-K., Downs, J., Dutta, R., ... Stewart, R. (2016). Cohort profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: current status and recent enhancement of an Electronic Mental Health Record-derived data resource. *BMJ Open*, 6(3), e008721. <https://doi.org/10.1136/bmjopen-2015-008721>
- Perkel, J. (2017a, 10). TechBlog: Interactive figures address data reproducibility. Consulté 29 octobre 2017, à l'adresse http://blogs.nature.com/naturejobs/2017/10/20/techblog-interactive-figures-address-data-reproducibility/?WT.ec_id=NEWSDAILY-20171027
- Perkel, J. (2017b, 09). TechBlog: Jupyter powers bioinformatics, again. Consulté 29 octobre 2017, à l'adresse <http://blogs.nature.com/naturejobs/2017/09/15/techblog-jupyter-powers-bioinformatics-again/>
- Perkins, D. O., Jeffries, C. D., Addington, J., Bearden, C. E., Cadenhead, K. S., Cannon, T. D., ... Heinssen, R. (2015). Towards a Psychosis Risk Blood Diagnostic for Persons Experiencing High-Risk Symptoms: Preliminary Results From the NAPLS Project. *Schizophrenia Bulletin*, 41(2), 419-428. <https://doi.org/10.1093/schbul/sbu099>
- Pettersson-Yeo, W., Benetti, S., Marquand, A. F., Dell'acqua, F., Williams, S. C., Allen, P., ... Mechelli, A. (2013). Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level., Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychological Medicine*, 43, 43(12, 12), 2547, 2547-2562. <https://doi.org/10.1017/S003329171300024X>, [10.1017/S003329171300024X](https://doi.org/10.1017/S003329171300024X)
- phatplus. (s. d.). 36 free vector icons of Analytics designed by phatplus. Consulté 13 novembre 2018, à l'adresse <https://www.flaticon.com/packs/analytics-7>
- Plus de droits pour vos données ! | CNIL. (s. d.). Consulté 14 novembre 2018, à l'adresse <https://www.cnil.fr/fr/plus-de-droits-pour-vos-donnees>
- Pon, D., & Coury, A. (2018). *Rapport final - Accélérer le virage numérique* (Stratégie de Transformation du Système de Santé). Consulté à l'adresse https://solidarites-sante.gouv.fr/IMG/pdf/masante2022_rapport_virage_numerique.pdf
- Popper, K. R. (1971). Conjectural knowledge: my solution of the problem of induction. *Revue Internationale de Philosophie*, 25(95/96 (1/2)), 167-197.
- Popper, K. R. (2002). *Conjectures and refutations: the growth of scientific knowledge*. London ; New York: Routledge.

- Post, F. H., Nielson, G. M., & Bonneau, G.-P. (Éd.). (2003). *Data visualization: the state of the art* (softcover reprint). New York, NY: Springer Science and Media.
- Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology*, 1-17. <https://doi.org/10.1007/s12553-017-0179-1>
- Pramparo, T., Pierce, K., Lombardo, M. V., Barnes, C. C., Marinero, S., Ahrens-Barbeau, C., ... Courchesne, E. (2015). Prediction of Autism by Translation and Immune/Inflammation Coexpressed Genes in Toddlers From Pediatric Community Practices. *JAMA Psychiatry*, 72(4), 386-394. <https://doi.org/10.1001/jamapsychiatry.2014.3008>
- Preoțiu-Pietro, D., Eichstaedt, J., Park, G., Sap, M., Smith, L., Tobolsky, V., ... Ungar, L. (2015). The role of personality, age, and gender in tweeting about mental illness. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality* (p. 21-30).
- Présidence de la République. (2018a, avril 4). Transcription du discours du Président de la République, Emmanuel Macron, sur l'intelligence artificielle. Consulté 17 septembre 2018, à l'adresse <http://www.elysee.fr/declarations/article/transcription-du-discours-du-president-de-la-republique-emmanuel-macron-sur-l-intelligence-artificielle/>
- Présidence de la République. (2018b, septembre 18). Transcription du discours sur la transformation du système de santé « Prendre soin de chacun » du Président de la République, Emmanuel Macron. Consulté 21 septembre 2018, à l'adresse <http://www.elysee.fr/declarations/article/transcription-du-discours-sur-la-transformation-du-systeme-de-sante-prendre-soin-de-chacun-du-president-de-la-republique-emmanuel-macron/>
- Product Platform Development & Pipeline - Mindstrong Health. (s. d.). Consulté 27 octobre 2018, à l'adresse <https://mindstronghealth.com/clinical-programs/>
- pubmeddev. (s. d.). Home - PubMed - NCBI. Consulté 2 novembre 2017, à l'adresse <https://www.ncbi.nlm.nih.gov/pubmed/>
- Pucheu, D. (2014). L'altérité à l'épreuve de l'ubiquité informationnelle. *Hermès, La Revue*, (68), 115-122.
- Qntfy Inc. (s. d.). Qntfy. Consulté 2 novembre 2017, à l'adresse <https://www.qntfy.com/>
- Quantified Self - Self Knowledge Through Numbers. (s. d.). Consulté 10 novembre 2018, à l'adresse <http://quantifiedself.com/>
- Qu'est-ce que le SNDS ? | SNDS. (s. d.). Consulté 31 octobre 2017, à l'adresse <https://www.snds.gouv.fr/SNDS/Qu-est-ce-que-le-SNDS>
- Redlich, R., Opel, N., Grotegerd, D., Dohm, K., Zaremba, D., Bürger, C., ... Dannlowski, U. (2016). Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data. *JAMA Psychiatry*, 73(6), 557-564. <https://doi.org/10.1001/jamapsychiatry.2016.0316>
- Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 | CNIL. (s. d.). Consulté 12 octobre 2017, à l'adresse <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>
- Researching for tomorrow. (s. d.). Consulté 8 septembre 2018, à l'adresse <https://deepmind.com/applied/deepmind-health/working-partners/health-research-tomorrow/>

Réseau Sentinelles > France > Le réseau Sentinelles. (s. d.). Consulté 20 novembre 2017, à l'adresse <https://websenti.u707.jussieu.fr/sentiweb/?page=presentation>

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016a). Model-Agnostic Interpretability of Machine Learning. *arXiv:1606.05386 [cs, stat]*. Consulté à l'adresse <http://arxiv.org/abs/1606.05386>

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016b). « Why Should I Trust You? »: Explaining the Predictions of Any Classifier. *arXiv:1602.04938 [cs, stat]*. Consulté à l'adresse <http://arxiv.org/abs/1602.04938>

Ricœur, P. (1988). L'identité narrative. *Esprit*, (No. 140/141 (7/8)), 295-304.

Robinson, E. B., Howrigan, D., Yang, J., Ripke, S., Anttila, V., Duncan, L. E., ... Neale, B. M. (2014). Response to 'Predicting the diagnosis of autism spectrum disorder using gene pathway analysis'. *Molecular Psychiatry*, 19(8), 860-861. <https://doi.org/10.1038/mp.2013.125>

Roess, A. (2017). The Promise, Growth, and Reality of Mobile Health — Another Data-free Zone. *New England Journal of Medicine*, 377(21), 2010-2011. <https://doi.org/10.1056/NEJMp1713180>

Roque, F. S., Jensen, P. B., Schmock, H., Dalgaard, M., Andreatta, M., Hansen, T., ... Brunak, S. (2011). Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts. *PLoS Computational Biology*, 7(8). <https://doi.org/10.1371/journal.pcbi.1002141>

Rosenfeld, A., Zemel, R., & Tsotsos, J. K. (2018). The Elephant in the Room. *arXiv:1808.03305 [cs]*. Consulté à l'adresse <http://arxiv.org/abs/1808.03305>

Roser, M., & Ritchie, H. (2017). Technological Progress. Consulté 25 octobre 2017, à l'adresse <https://ourworldindata.org/technological-progress/>

Ross, C., & Swetlitz, I. (2017, septembre 5). IBM pitched Watson as a revolution in cancer care. It's nowhere close. Consulté 9 octobre 2017, à l'adresse <https://www.statnews.com/2017/09/05/watson-ibm-cancer/>

Ross, C., & Swetlitz, I. (2018, juillet 25). IBM's Watson recommended « unsafe and incorrect » cancer treatments. Consulté 31 octobre 2018, à l'adresse <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/>

Rothstein, M. A. (2010). Is Deidentification Sufficient to Protect Health Privacy in Research? *The American Journal of Bioethics*, 10(9), 3-11. <https://doi.org/10.1080/15265161.2010.494215>

Rouvroy, A., & Berns, T. (2010). Le nouveau pouvoir statistique. *Multitudes*, (40), 88-103. <https://doi.org/10.3917/mult.040.0088>

SAIL Databank - The Secure Anonymised Information Linkage Databank. (s. d.). Consulté à l'adresse <https://saildatabank.com/about-us/overview/>

Schlosser, D. A., Campellone, T. R., Truong, B., Etter, K., Vergani, S., Komaiko, K., & Vinogradov, S. (2018). Efficacy of PRIME, a Mobile App Intervention Designed to Improve Motivation in Young People With Schizophrenia. *Schizophrenia Bulletin*, 44(5), 1010-1020. <https://doi.org/10.1093/schbul/sby078>

Schmaal, L., Marquand, A. F., Rhebergen, D., van Tol, M.-J., Ruhé, H. G., van der Wee, N. J. A., ... Penninx, B. W. J. H. (2015). Predicting the Naturalistic Course of Major Depressive Disorder Using Clinical and Multimodal Neuroimaging Information: A Multivariate Pattern Recognition Study. *Biological Psychiatry*, 78(4), 278-286. <https://doi.org/10.1016/j.biopsych.2014.11.018>

- Schmidt, A., Cappucciati, M., Radua, J., Rutigliano, G., Rocchetti, M., Dell'Osso, L., ... Fusar-Poli, P. (2017). Improving Prognostic Accuracy in Subjects at Clinical High Risk for Psychosis: Systematic Review of Predictive Models and Meta-analytical Sequential Testing Simulation. *Schizophrenia Bulletin*, 43(2), 375-388. <https://doi.org/10.1093/schbul/sbw098>
- Schnack, H. G., Nieuwenhuis, M., van Haren, N. E. M., Abramovic, L., Scheewe, T. W., Brouwer, R. M., ... Kahn, R. S. (2014). Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *NeuroImage*, 84(Supplement C), 299-306. <https://doi.org/10.1016/j.neuroimage.2013.08.053>
- Schoenbaum, M., Kessler, R. C., Gilman, S. E., Colpe, L. J., Heeringa, S. G., Stein, M. B., ... Cox, K. L. (2014). Predictors of Suicide and Accident Death in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS): Results From the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *JAMA Psychiatry*, 71(5), 493-503. <https://doi.org/10.1001/jamapsychiatry.2013.4417>
- Schück, S., Foulquié, P., Mebarki, A., Voillot, P., Texier, N., & Faviez, C. (2018). Détection automatique du mésusage des neuroleptiques dans le trouble anxieux et la démence à partir des réseaux sociaux. *Revue d'Épidémiologie et de Santé Publique*, 66, S211-S212. <https://doi.org/10.1016/j.respe.2018.04.048>
- Science – Medibio. (s. d.). Consulté 26 octobre 2018, à l'adresse <https://medibio.com.au/science/>
- Seyffert, M., Lagisetty, P., Landgraf, J., Chopra, V., Pfeiffer, P. N., Conte, M. L., & Rogers, M. A. M. (2016). Internet-Delivered Cognitive Behavioral Therapy to Treat Insomnia: A Systematic Review and Meta-Analysis. *PLoS ONE*, 11(2). <https://doi.org/10.1371/journal.pone.0149139>
- Sheridan, K. (2018a, septembre 13). Here's the data behind the new Apple Watch EKG app. Consulté 14 septembre 2018, à l'adresse <https://www.statnews.com/2018/09/13/heres-the-data-behind-the-new-apple-watch-ekg-app/>
- Sheridan, K. (2018b, octobre 10). AI for drug development: What's possible and what's just hype? Consulté 18 octobre 2018, à l'adresse <https://www.statnews.com/2018/10/10/ai-big-data-panel/>
- Siegle, G. J., Thompson, W. K., Collier, A., Berman, S. R., Feldmiller, J., Thase, M. E., & Friedman, E. S. (2012). Toward Clinically Useful Neuroimaging in Depression Treatment: Prognostic Utility of Subgenual Cingulate Activity for Determining Depression Outcome in Cognitive Therapy Across Studies, Scanners, and Patient Characteristics. *Archives of General Psychiatry*, 69(9), 913-924. <https://doi.org/10.1001/archgenpsychiatry.2012.65>
- Silverman, E. (2018, juin 28). Amazon agrees to buy PillPack in a bid to remake the pharma supply chain. Consulté 9 septembre 2018, à l'adresse <https://www.statnews.com/pharmalot/2018/06/28/amazon-pillpack-distribution-pharmaceuticals/>
- Simon, G. E., Johnson, E., Lawrence, J. M., Rossom, R. C., Ahmedani, B., Lynch, F. L., ... Shortreed, S. M. (2018). Predicting Suicide Attempts and Suicide Deaths Following Outpatient Visits Using Electronic Health Records. *American Journal of Psychiatry*, 175(10), 951-960. <https://doi.org/10.1176/appi.ajp.2018.17101167>
- Singh, S., Okun, A., & Jackson, A. (2017). Artificial intelligence: Learning to play Go from scratch. *Nature*, 550(7676), 336-337. <https://doi.org/10.1038/550336a>
- Skafidas, E., Testa, R., Zantomio, D., Chana, G., Everall, I. P., & Pantelis, C. (2014a). Predicting the diagnosis of autism spectrum disorder using gene pathway analysis., Predicting the diagnosis

- of autism spectrum disorder using gene pathway analysis. *Molecular Psychiatry, Molecular Psychiatry*, 19, 19(4, 4), 504, 504-510. <https://doi.org/10.1038/mp.2012.126>, 10.1038/mp.2012.126
- Skafidas, E., Testa, R., Zantomio, D., Chana, G., Everall, I. P., & Pantelis, C. (2014b). Response to Belgard et al. *Molecular Psychiatry*, 19(4), 407-409. <https://doi.org/10.1038/mp.2013.186>
- Skafidas, E., Testa, R., Zantomio, D., Chana, G., Everall, I. P., & Pantelis, C. (2015). Response to Robinson et al. *Molecular Psychiatry*, 20(7), 794-794. <https://doi.org/10.1038/mp.2015.15>
- Stack Overflow - Where Developers Learn, Share, & Build Careers. (s. d.). Consulté 19 octobre 2018, à l'adresse <https://stackoverflow.com/>
- Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., ... Robinson, G. E. (2015). Big Data: Astronomical or Genomical? *PLOS Biology*, 13(7), e1002195. <https://doi.org/10.1371/journal.pbio.1002195>
- Stern, S., Linker, S., Vadodaria, K. C., Marchetto, M. C., & Gage, F. H. (2018). Prediction of response to drug therapy in psychiatric disorders. *Open Biology*, 8(5). <https://doi.org/10.1098/rsob.180031>
- Suffin, S. C., Emory, W. H., D, N. G. M., Arora, G. S., Schiller, M. J., & D, A. K. M. (2007). A QEEG Database Method for Predicting Pharmacotherapeutic Outcome in Refractory Major Depressive Disorders. *Journal of American Physicians and Surgeons*, 12(4), 104-108.
- Sullivan, P. F. (2010). The Psychiatric GWAS Consortium: Big Science Comes to Psychiatry. *Neuron*, 68(2), 182-186. <https://doi.org/10.1016/j.neuron.2010.10.003>
- Swetlitz, I., & Ross, C. (2018, octobre 19). Head of IBM Watson Health leaving post after company stumbles, growing criticism. Consulté 31 octobre 2018, à l'adresse <https://www.statnews.com/2018/10/19/head-of-ibm-watson-health-leaving-post/>
- Talkspace. (s. d.). Talkspace | Online Therapy | Counseling Online | Marriage Counseling. Consulté 26 octobre 2018, à l'adresse <https://www.talkspace.com>
- Tansey, K. E., Guipponi, M., Perroud, N., Bondolfi, G., Domenici, E., Evans, D., ... Uher, R. (2012). Genetic predictors of response to serotonergic and noradrenergic antidepressants in major depressive disorder: a genome-wide analysis of individual-level data and a meta-analysis. *PLoS Medicine*, 9(10), e1001326. <https://doi.org/10.1371/journal.pmed.1001326>
- Taylor, A. S., & Vincent, J. (2005). An SMS History. In C. Sanger, L. Hamill, A. Lasen, & D. Diaper (Éd.), *Mobile World: Past, Present and Future* (p. 75-91). London: Springer London. https://doi.org/10.1007/1-84628-204-7_5
- Technology - Fitbit. (s. d.). Consulté 19 novembre 2017, à l'adresse <https://www.fitbit.com/technology>
- Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, 10(7), 309-318. <https://doi.org/10.1016/j.tics.2006.05.009>
- The Wachowski Brothers. (1999). *The Matrix*. Consulté à l'adresse <http://www.imdb.com/title/tt0133093/>
- Thompson, N. (2018, mars 31). Emmanuel Macron Talks to WIRED About France's AI Strategy. Consulté 18 avril 2018, à l'adresse <https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/>

- Torous, J., & Baker, J. T. (2016). Why Psychiatry Needs Data Science and Data Science Needs Psychiatry: Connecting With Technology. *JAMA Psychiatry*, 73(1), 3. <https://doi.org/10.1001/jamapsychiatry.2015.2622>
- Torous, J., Chan, S. R., Tan, S. Y.-M., Behrens, J., Mathew, I., Conrad, E. J., ... Keshavan, M. (2014). Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. *JMIR Mental Health*, 1(1), e5. <https://doi.org/10.2196/mental.4004>
- Uher, R., Perlis, R. H., Henigsberg, N., Zobel, A., Rietschel, M., Mors, O., ... McGuffin, P. (2012). Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms., Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. *Psychological Medicine*, 42, 42(5, 5), 967, 967-980. <https://doi.org/10.1017/S0033291711001905>, [10.1017/S0033291711001905](https://doi.org/10.1017/S0033291711001905)
- United Nations. (1999). *The World at Six Billion*. Consulté à l'adresse <http://www.un.org/esa/population/publications/sixbillion/sixbillion.htm>
- van Loo, H. M., Wanders, R. B. K., Wardenaar, K. J., & Fried, E. I. (2018). Problems with latent class analysis to detect data-driven subtypes of depression. *Molecular Psychiatry*, 23(3), 495-496. <https://doi.org/10.1038/mp.2016.202>
- Verger, P., Mmadi Mrenda, B., Cortaredona, S., Tournier, M., & Verdoux, H. (2018). Dynamiques temporelles de remboursements de médicaments anxiolytiques chez des nouveaux consommateurs âgés de 50 ans et plus. *Revue d'Épidémiologie et de Santé Publique*, 66, S217-S218. <https://doi.org/10.1016/j.respe.2018.04.014>
- Verily Life Sciences. (s. d.). Consulté 30 octobre 2017, à l'adresse <https://verily.com/projects/>
- Villani, C. (2018). *Donner un sens à l'intelligence artificielle - Pour une stratégie nationale et européenne* (Mission parlementaire du 8 septembre 2017 au 8 mars 2018).
- Vos, T., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., ... Murray, C. J. L. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, 390(10100), 1211-1259. [https://doi.org/10.1016/S0140-6736\(17\)32154-2](https://doi.org/10.1016/S0140-6736(17)32154-2)
- Wang, R., Blackburn, G., Desai, M., Phelan, D., Gillinov, L., Houghtaling, P., & Gillinov, M. (2017). Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA Cardiology*, 2(1), 104-106. <https://doi.org/10.1001/jamacardio.2016.3340>
- Watson, G. W., & Calhoun, E. C. (1960). Brief History of Computers. *School Science and Mathematics*, 60(2), 87-94. <https://doi.org/10.1111/j.1949-8594.1960.tb08329.x>
- Weber, G. M., Mandl, K. D., & Kohane, I. S. (2014). Finding the Missing Link for Big Biomedical Data. *JAMA*, 311(24), 2479-2480. <https://doi.org/10.1001/jama.2014.4228>
- Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*, 265(3), 94-104. <https://doi.org/10.1038/scientificamerican0991-94>
- Wensing, T., Cieslik, E. C., Müller, V. I., Hoffstaedter, F., Eickhoff, S. B., & Nickl-Jockschat, T. (2017). Neural correlates of formal thought disorder: An activation likelihood estimation meta-analysis. *Human Brain Mapping*, 38(10), 4946-4965. <https://doi.org/10.1002/hbm.23706>

- Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., ... Garavan, H. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*, 512(7513), 185-189. <https://doi.org/10.1038/nature13402>
- Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., ... Ziesch, V. (2014). Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*, 512(7513), 185-189. <https://doi.org/10.1038/nature13402>
- Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., ... Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, 21(5), 680-685. <https://doi.org/10.1038/mp.2015.109>
- WHO | ICTRP International Clinical Trials Registry Platform. (s. d.). Consulté 12 février 2018, à l'adresse <http://www.who.int/ictcp/en/>
- Woebot - Your charming robot friend who is ready to listen, 24/7. (s. d.). Consulté 25 juillet 2018, à l'adresse <https://www.woebot.io/>
- Wolf, G. (2010). *The quantified self*. TED@Cannes. Consulté à l'adresse https://www.ted.com/talks/gary_wolf_the_quantified_self
- Woo, C.-W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: brain models in translational neuroimaging. *Nature Neuroscience*, 20(3), 365-377. <https://doi.org/10.1038/nn.4478>
- World Health Organization (Éd.). (2009). *Public health campaigns: getting the message across*. Geneva, Switzerland: World Health Organization.
- World Population Growth. (s. d.). Consulté 25 octobre 2017, à l'adresse <https://ourworldindata.org/world-population-growth/>
- Wu, M.-J., Mwangi, B., Bauer, I. E., Passos, I. C., Sanches, M., Zunta-Soares, G. B., ... Soares, J. C. (2017). Identification and individualized prediction of clinical phenotypes in bipolar disorders using neurocognitive data, neuroimaging scans and machine learning. *NeuroImage*, 145, 254-264. <https://doi.org/10.1016/j.neuroimage.2016.02.016>
- Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K. C., ... Frey, B. J. (2015). The human splicing code reveals new insights into the genetic determinants of disease. *Science*, 347(6218), 1254806. <https://doi.org/10.1126/science.1254806>
- xkcd: Move Fast and Break Things. (s. d.). Consulté 31 octobre 2018, à l'adresse <https://xkcd.com/1428/>
- Yang, H., Liu, J., Sui, J., Pearlson, G., & Calhoun, V. D. (2010). A Hybrid Machine Learning Method for Fusing fMRI and Genetic Data: Combining both Improves Classification of Schizophrenia. *Frontiers in Human Neuroscience*, 4. <https://doi.org/10.3389/fnhum.2010.00192>
- youtube.com/AlfredoCreates. (s. d.). « Social Media FREE » by youtube.com/AlfredoCreates. Consulté 13 novembre 2018, à l'adresse <https://www.iconfinder.com/iconsets/social-media-free-13>
- Zarogianni, E., Moorhead, T. W. J., & Lawrie, S. M. (2013). Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. *NeuroImage: Clinical*, 3(Supplement C), 279-289. <https://doi.org/10.1016/j.nicl.2013.09.003>
- Zenodo - Research. Shared. (s. d.). Consulté 24 septembre 2017, à l'adresse <http://about.zenodo.org/policies/>

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist — it's time to make it fair. *Nature*, 559(7714), 324. <https://doi.org/10.1038/d41586-018-05707-8>

Zweig, M., Tecco, H., & Huang, M. (2018). 2018 Midyear Funding Review: Digital health déjà vu in yet another record breaking half. Consulté 18 octobre 2018, à l'adresse <https://rockhealth.com/reports/2018-midyear-funding-review-digital-health-deja-vu-in-yet-another-record-breaking-half/>

Annexe 1 : Synthèse des études en revue

Découvrir		
Publication	Pathologie	Modalité étudiée
(Hazlett et al., 2017)	Autisme Haut-risque d'autisme	IRM morphologique
(Just et al., 2014)	Autisme	IRMf
(Xiong et al., 2015)	Autisme	Épissage ARNm
(De Choudhury et al., 2013)	Dépression	Réseaux sociaux internet
(Preoțiu-Pietro et al., 2015)	Dépression Syndrome de stress post-traumatique	Twitter
(Wensing et al., 2017)	Dissociation psychique	Imagerie fonctionnelle BrainMap
(De Choudhury et al., 2016)	Idéation suicidaire	Réseaux sociaux internet
(Just et al., 2017)	Idéation suicidaire	IRMf
(Bedi et al., 2015)	Haut-risque psychotique	Analyse du discours libre
(Ben-Zeev et al., 2017)	Schizophrénie	Smartphone : marqueurs comportementaux, auto-évaluation
(Brodersen et al., 2014)	Schizophrénie	IRMf : dynamic causal modeling
(Natalia B. Mota et al., 2012)	Schizophrénie Trouble bipolaire	Analyse quantitative du discours
(Natália B. Mota et al., 2014)	Schizophrénie Trouble bipolaire	Analyse quantitative du discours
(Wu et al., 2017)	Trouble bipolaire	IRM de diffusion tests neurocognitifs
(Gillan & Whelan, 2017)	Transdiagnostic	Tâche cognitive en ligne Questionnaires d'auto-évaluation en ligne

Classifier		
Dossiers médicaux électroniques	Publication	Pathologie
	(Castro et al., 2015)	Trouble bipolaire
	(Jackson et al., 2017)	Trouble mental sévère
	(Roque et al., 2011)	Multiplés
Imagerie	Publication	Pathologie
	(Drysdale et al., 2017)	Dépression
	(Y. J. Chen et al., 2018)	Schizophrénie
	(Kawasaki et al., 2007)	Schizophrénie
	(Nieuwenhuis et al., 2012)	Schizophrénie
	(Schnack et al., 2014)	Schizophrénie Trouble bipolaire
	(Librenza-Garcia et al., 2017)	Trouble bipolaire
	(Mwangi et al., 2016)	Trouble bipolaire
Génétique	Publication	Pathologie
	(CHARGE inflammation working group et al., 2016)	Dépression
	(Chuang & Kuo, 2017)	Trouble bipolaire
	(Maier et al., 2015)	Dépression Schizophrénie Trouble bipolaire
	(Skafidas et al., 2014a)	Autisme
	(Yang et al., 2010)	Schizophrénie
Biomarqueurs sériques	Publication	Pathologie
	(Chan et al., 2015)	Schizophrénie
	(Haenisch et al., 2016)	Trouble bipolaire
Multimodal	Publication	Pathologie
	(Clementz et al., 2015)	Psychose transdiagnostique
	(Grisanzio et al., 2018)	Transdiagnostic : humeur, anxiété, trauma
	(Whelan, Watts, Orr, Althoff, Artiges, Banaschewski, Barker, Ziesch, et al., 2014)	Mésusage d'alcool

Prédire : prévention		
Publication	Pathologie	Modalité étudiée
(Glatt et al., 2012)	Autisme	Transcriptome sang périphérique
(Hazlett et al., 2017)	Autisme	IRM morphologique
(Pramparo et al., 2015)	Autisme	Transcriptome périphérique
(Farb et al., 2011)	Dépression	IRMf
(Schmaal et al., 2015)	Dépression	Clinique Imagerie multimodale
(Bedi et al., 2015)	Haut-risque psychotique	Analyse automatisée du discours
(Cannon et al., 2016)	Haut-risque psychotique	Score de risque sur questionnaire clinique
(Carrión et al., 2016)	Haut-risque psychotique	Score de risque sur questionnaire clinique (validation externe de (Cannon et al., 2016)
(Ciarleglio et al., 2018)	Haut-risque psychotique	Questionnaire clinique
(Koutsouleris et al., 2015)	Haut-risque psychotique	IRM morphologique
(Koutsouleris, Borgwardt, et al., 2012)	Haut-risque psychotique	IRM morphologique
(Koutsouleris, Davatzikos, et al., 2012)	Haut-risque psychotique	IRMf
(Nieman et al., 2014)	Haut-risque psychotique	Électrophysiologie, Clinique
(Perkins et al., 2015)	Haut-risque psychotique	Biomarqueurs sériques
(Pettersson-Yeo et al., 2013)	Haut-risque psychotique	Génomique Neurocognitif Imagerie multimodale
(Schmidt et al., 2017)	Haut-risque psychotique	Méta-analyse
(Whelan, Watts, Orr, Althoff, Artiges, Banaschewski, Barker, Bokde, et al., 2014)	Mésusage d'alcool	IRM Personnalité Fonctionnement cognitif Environnement Génétique Autobiographie
(Chan et al., 2015)	Schizophrénie	Biomarqueurs sériques
(Kessler et al., 2015)	Suicide	Dossiers administratifs et médicaux de l'armée : données démographiques, sociales, professionnelles, juridiques, médicales
(McCoy et al., 2016)	Suicide	Dossiers médicaux électroniques
(Niculescu et al., 2015)	Suicide	Génomique, Clinique
(Simon et al., 2018)	Suicide	Dossiers médicaux électroniques

Prédire : optimisation du traitement			
Publication	Pathologie	Traitement	Modalité étudiée
(Whitfield-Gabrieli et al., 2016)	Anxiété sociale	TCC	IRMf de repos IRM de diffusion
(Adam M. Chekroud et al., 2017)	Dépression	Antidépresseurs multiples	Essais cliniques historiques
(Adam Mourad Chekroud et al., 2016)	Dépression	Antidépresseurs multiples	Essais cliniques historiques
(DeBattista et al., 2011)	Dépression	Antidépresseurs multiples	EEG
(Drysdale et al., 2017)	Dépression	rTMS	IRMf de repos
(Etkin et al., 2015)	Dépression	Antidépresseurs multiples	Fonctionnement cognitivo-émotionnel
(Redlich et al., 2016)	Dépression	ECT	IRM morphologique
(Siegle et al., 2012)	Dépression	TCC	IRMf
(Tansey et al., 2012)	Dépression	Antidépresseurs multiples	Génomique
(Uher et al., 2012)	Dépression	Antidépresseurs multiples	Essais cliniques historiques
(Koutsouleris et al., 2016)	Premier épisode psychotique	Antipsychotiques multiples	Questionnaires cliniques Fonctionnement neurocognitif Socio-démographiques
(International Consortium on Lithium Genetics (ConLi+Gen) et al., 2018)	Trouble bipolaire	Lithium	Génomique

Suivre		
Publication	Pathologie	Intervention
(Gustafson et al., 2014)	Dépendance à l'alcool	A-CHESS : soutien en post-réhabilitation par des fonctionnalités statiques (ressources d'informations), interactives (auto-évaluation, forums...), dynamiques (bouton secours)
(Dagum, 2018)	Non précisé	Surveillance du fonctionnement cognitif par marqueurs numériques (smartphone)
(Schück et al., 2018)	Pharmacovigilance : mésusage des antipsychotiques	Analyse textuelle des réseaux sociaux
(Foulquié et al., 2018)	Pharmacovigilance : usages du méthylphénidate	Analyse textuelle des réseaux sociaux
(Fitzpatrick et al., 2017)	Population infra-cliniques (symptômes anxio-dépressifs)	Woebot : Agent conversationnel automatique offrant un soutien basé sur les principes de TCC
(Ben-Zeev et al., 2017)	Schizophrénie	CrossCheck : marqueurs comportementaux recueillis par smartphone, questionnaires d'auto-évaluation rythmés
(Bucci et al., 2018)	Schizophrénie	Actissist : questionnaires d'auto-évaluation rythmés, conseils automatiques en fonction des problèmes relevés
(Schlosser et al., 2018)	Schizophrénie	PRIME : soutien motivationnel, interactions avec des pairs
(Coppersmith et al., 2018)	Suicide	Dépistage du risque suicidaire sur réseaux sociaux
(Verger et al., 2018)	Suivi pharmacologique	Analyse dynamique des données de remboursement

Annexe 2 : Matrice RDoC

Adaptée de (« NIMH » RDoC Matrix », s. d.)

Negative Valence Systems

Construct/ Subconstruct	Genes	Molecules	Cells	Circuits	Physiology	Behavior	Self-Report	Paradigms
Acute Threat ("Fear")								
Potential Threat ("Anxiety")								
Sustained Threat								
Loss								
Frustrative Nonreward								

Positive Valence Systems

Construct/Subconstruct		Genes (...)	Paradigms
Reward Responsiveness	Reward Anticipation		
	Initial Response to Reward		
	Reward Satiation		
Reward Learning	Probabilistic and Reinforcement Learning		
	Reward Prediction Error		
	Habit		
Reward Valuation	Reward (probability)		
	Delay		
	Effort		

Cognitive Systems

Construct/Subconstruct		Genes (...)	Paradigms
Attention			
Perception	Visual Perception		
	Auditory Perception		
	Olfactory/Somatosensory/Multimodal/Perception		
Declarative Memory			
Language			
Cognitive Control	Goal Selection; Updating, Representation, and Maintenance		
	Response Selection; Inhibition/Suppression		
	Performance Monitoring		
Working Memory	Active Maintenance		
	Flexible Updating		
	Limited Capacity		
	Interference Control		

Social Processes

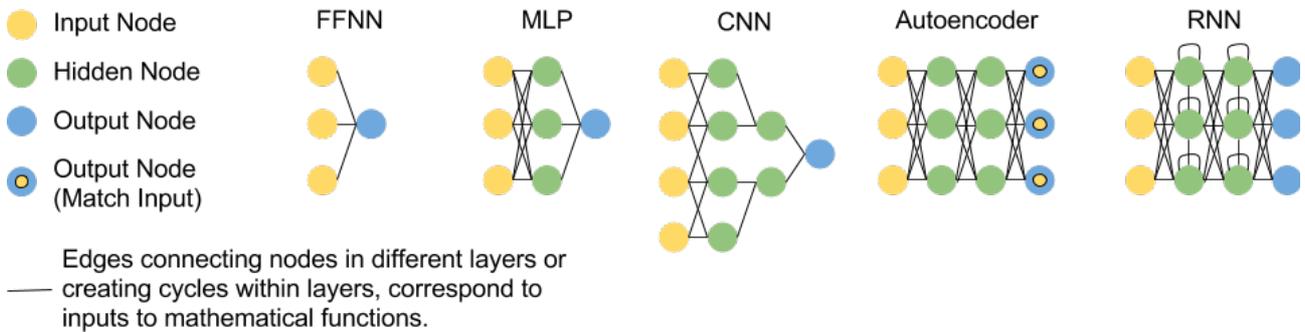
Construct/Subconstruct		Genes (...)	Paradigms
Affiliation and Attachment			
Social Communication	Reception of Facial Communication		
	Production of Facial Communication		
	Reception of Non-Facial Communication		
	Production of Non-Facial Communication		
Perception and Understanding of Self	Agency		
	Self-Knowledge		
Perception and Understanding of Others	Animacy Perception		
	Action Perception		
	Understanding Mental States		

Arousal and Regulatory Systems

Construct/Subconstruct	Genes	(...)	Paradigms
Arousal			
Circadian Rhythms			
Sleep-Wakefulness			

Annexe 3 : Architectures de réseaux de neurones artificiels pour apprentissage profond

Adapté de (Ching et al., 2018)



Nom	Définition	Exemples d'utilisation
<i>Feed-forward neural network (FFNN)</i>	Réseau de neurones sans liens cycliques entre nœuds d'une même couche	Utilisation fréquente et variée
<i>Multi-layer perceptron (MLP)</i>	FFNN avec au moins une couche cachée, où chaque couche suivante est une fonction non-linéaire de la précédente	N'imposent pas de structure Sont fréquemment utilisés quand les <i>inputs</i> ne sont pas naturellement ordonnés (ex : mesures d'expression génique)
<i>Convolutional neural network (CNN)</i>	Réseau avec des couches où la connectivité entre nœuds est plus localisée. Souvent une bonne performance, si utilisées sur des données adaptées. Peuvent être entraînés sur un échantillon plus petit car ont moins de paramètres, avec une meilleure efficacité	Utilisés pour données séquentielles (séquences ADN par exemple) ou données en grille (imagerie médicale ou de microscopie)
<i>Recurrent neural network (RNN)</i>	Réseau de neurones avec cycles entre nœuds dans une couche cachée	Utilisée pour données séquentielles comme : texte, séquences génomiques, séries temporelles cliniques

Nom	Définition	Exemples d'utilisation
<i>Long short-term memory (LSTM) neural network</i>	Réseau possédant des caractéristiques lui permettant de capturer des dépendances à plus long terme entre les variables	Leur utilisation se propage rapidement dans l'analyse du langage naturel. Seront probablement plus largement utilisés pour les données biologiques séquentielles
<i>Autoencoder (AE)</i>	Réseau où l'objectif d'entraînement est la minimisation de l'erreur entre la couche <i>output</i> et la couche <i>input</i> . Non-supervisés, souvent utilisés pour la réduction de dimensions.	Exemples d'utilisation pour l'analyse non-supervisée de données d'expression génique, et de données le dossier médicaux électroniques
<i>Variational autoencoder (VAE)</i>	Autoencodeur avec la contrainte supplémentaire que le modèle est entraîné pour apprendre des caractéristiques à distribution normale	Exemples répétés de leur efficacité pour produire des représentations réduites en imagerie. Quelques publications sur leur utilisation dans l'analyse d'expression génique.
<i>Denoising autoencoder (DA)</i>	Autoencodeur avec une étape supplémentaire où du bruit est ajouté au <i>input</i> pendant l'entraînement. L'étape de débruitage permet de lisser les données, et peut permettre l'utilisation efficace sur des données avec beaucoup de bruit.	Utilisés comme les AE dans l'analyse non-supervisée de données d'expression génique et de dossiers médicaux
<i>Generative neural network</i>	Peuvent générer des données semblables à des données <i>input</i> . Peuvent être utilisés pour produire des exemples hypothétiques	Certains des réseaux non-supervisés cités plus haut peuvent être utilisés à cette fin
<i>Restricted Boltzmann machine (RBM)</i>	Réseau génératif qui est une des bases de nombreuses approches d'apprentissage profond : une seule couche <i>input</i> , une seule couche cachée, sans connexions entre les nœuds à l'intérieur de chaque couche	Ont été utilisés pour combiner plusieurs types de données « omiques » (méthylation d'ADN, transcriptome, expression d'ARN interférent, par exemple)
<i>Deep belief network (DBN)</i>	Réseau génératif avec plusieurs couches cachées, pouvant être obtenu en combinant plusieurs RBM	Peuvent être utilisés pour prédire de nouveaux liens dans un réseau d'interaction entre une molécule (médicament par exemple) et une cible

Nom	Définition	Exemples d'utilisation
<i>Generative adversarial network (GAN)</i>	Approche générative où deux réseaux sont entraînés. L'un, le générateur, est fourni avec un jeu de <i>inputs</i> générés au hasard, et doit en créer des échantillons. Le deuxième, le discriminateur, est entraîné à distinguer entre échantillons réels et générés. Après leur entraînement l'un contre l'autre, le générateur résultant peut être utilisé pour créer de nouveaux échantillons/exemples.	Peuvent créer de nouveaux exemples de données ayant les mêmes propriétés statistiques que des jeux de données identifiables et qui ne sont pas autorisés à être partagés (afin de partager ces données synthétiques et non les originales). Aussi utilisés pour générer des images de microscopie

« Par délibération de son Conseil en date du 10 Novembre 1972, l'Université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ou mémoires. Ces opinions doivent être considérées comme propres à leurs auteurs ».

VU, le Président de Thèse

VU, le Doyen de la Faculté

VU et permis d'imprimer
en référence à la délibération
du Conseil d'Université
en date du 14 Décembre 1973

Pour le Président
de l'Université de CAEN et P.O

Le Doyen

ANNEE DE SOUTENANCE : 2018

NOM ET PRENOM DE L'AUTEUR : BELHADI Aisha

Utilisations actuelles et impact du Big Data en Psychiatrie

RESUME DE LA THESE EN FRANÇAIS : Le succès du traitement algorithmique des données massives dans le domaine commercial suscite l'engouement des pouvoirs publics, du secteur privé et de la recherche biomédicale pour un déploiement massif des technologies du Big Data en santé. La promesse d'une transformation du soin par l'exploitation des données afin de réaliser des prédictions individuelles rejoint le discours autour d'une médecine prédictive, préventive, personnalisée et participative. Ce travail s'attache dans un premier temps à définir le Big Data en décrivant ses techniques et son architecture, et en retraçant son contexte politique et culturel, notamment dans le domaine de la santé. Cette recontextualisation permet, dans un deuxième temps, d'évaluer l'impact du Big Data sur la pratique clinique en psychiatrie. Une revue sélective de la littérature médicale et scientifique identifie d'abord des exemples d'utilisation du Big Data dans la découverte scientifique, et dans la classification, la prédiction et le suivi de pathologies psychiatriques. Une critique méthodologique de ces études montre que l'utilisation du Big Data en psychiatrie est encore au stade préliminaire, loin de la révolution des pratiques annoncée, et que la complexité méthodologique de la recherche en psychiatrie ne pourra pas être contournée de sitôt. Ensuite, partant de différentes grilles de lecture, une analyse critique du Big Data en psychiatrie montre que cette technologie soulève des interrogations autour de l'éthique, de l'épistémologie et de la subjectivité. Une réflexion sur ces interrogations sous l'angle de la clinique est proposée ici, avec une incitation à l'approfondir davantage.

MOTS CLES : Psychiatrie ; Big Data ; apprentissage automatique ; intelligence artificielle ; médecine de précision ; santé mentale ; épistémologie.

Current uses and impact of Big Data in Psychiatry

RESUME DE LA THESE EN ANGLAIS : The success of Big Data analytics and machine learning in business and the technology industry is driving an enthusiastic push by public authorities, the private sector and biomedical research for a massive deployment of Big Data in health. The promise of transforming healthcare through data-driven individual predictions is central to the discourse surrounding precision medicine, also described as predictive, preventive, personalized and participative. The current work begins with a definition of Big Data, based on a description of its techniques and architecture, as well as its political and cultural contexts, with a focus on healthcare. This recontextualization allows us to then evaluate the impact of Big Data on clinical practice in psychiatry. First, through a selective review of the medical and scientific literature, examples of uses of Big Data in psychiatry are identified. They pertain to the classification, prediction and monitoring of psychiatric disorders, as well as to scientific discovery in psychiatry. A critical analysis of their statistical methodology reveals that the use of Big Data in psychiatry is still at a preliminary stage, quite far from affecting real change in clinical practice, and that the methodological complexity of psychiatric research will not be solved anytime soon. Next, through a number of theoretical critical perspectives, we show that Big Data in psychiatry raises questions related to ethics, epistemology and subjectivity/personhood. A clinical reflection on these issues is proposed here, with a call for further clinical readings of the questions raised by Big Data in psychiatry.

KEY WORDS : Psychiatry ; Big Data ; machine learning ; artificial intelligence ; precision medicine ; mental health ; epistemology.