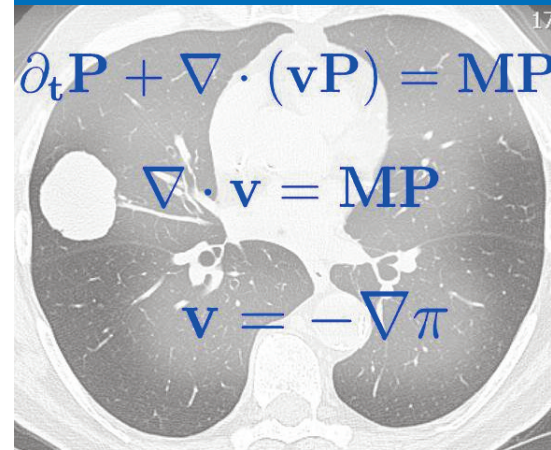




INSTITUT DE FRANCE
Académie des sciences



Mathématiques, mégadonnées et santé : l'exemple du cancer

28 novembre 2017 de 14h30 à 17h00

Grande salle des séances de
l'Institut de France

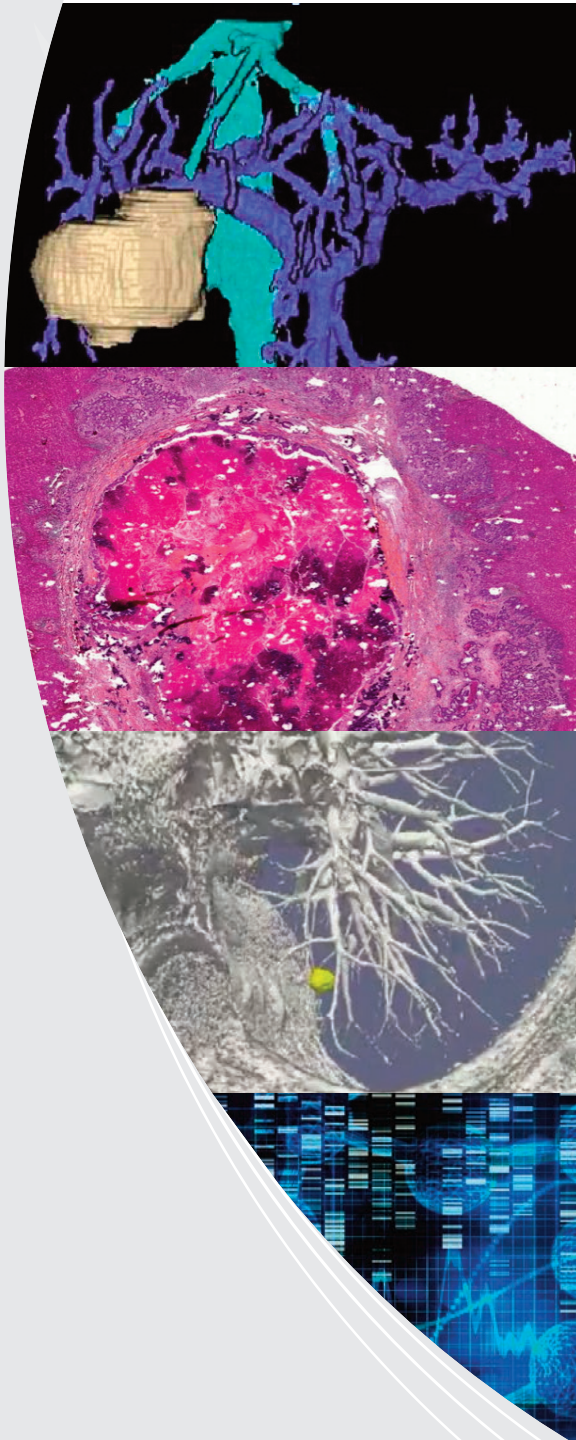
23, quai de Conti, 75006 Paris

La quantité des données collectées explose et l'analyse de ces mégadonnées ouvre d'importantes perspectives dans le domaine de la santé, notamment en cancérologie. De nombreux acteurs publics et privés, dont tous les géants du web, s'y investissent. Les grandes entreprises du secteur sont américaines ou chinoises. Cependant, la France dispose des talents pour être un acteur majeur dans ce domaine, grâce à des mathématiciens, des statisticiens, des physiciens de grande valeur et grâce à une recherche en cancérologie de haut niveau. Les académies de médecine et des sciences se sont associées pour créer un groupe de travail dédié afin de stimuler la réflexion, mettre en place des coopérations et faire des propositions pour développer cette recherche.

Plusieurs facettes seront explorées. Le SNIIRAM (Système national d'information inter-régime de l'Assurance maladie), véritable trésor d'informations, ouvre l'accès, pour la recherche, aux mégadonnées publiques, sous réserve de concilier besoins scientifiques et protection de la vie privée.

Les nouvelles méthodologies permettent de modéliser la progression des tumeurs en fonction du traitement grâce à des algorithmes qui prennent en compte les données d'imagerie mais aussi le génome des cellules.

La recherche en cancérologie va sans doute changer de paradigme : la séquence - hypothèse préclinique, études cliniques phase I, II, III - laissera la place aux plateformes dont les grandes quantités de données permettent de croiser phénotype et génotype. Les études ciblent désormais les anomalies génétiques plus que les organes.



Les organisateurs du colloque



Nicholas AYACHE

Nicholas Ayache est directeur de recherche de classe exceptionnelle à l'Institut national de recherche en informatique et en automatique (Inria), et membre de l'Académie des sciences. Il a été professeur invité au Collège de France, titulaire de la chaire annuelle Informatique et sciences numériques (2013-2014), où il a introduit un nouveau cours sur le « patient numérique personnalisé ». Ses recherches actuelles se concentrent sur l'analyse informatique des images médicales et sur le développement de la « médecine numérique ».



Bernard NORDLINGER

Bernard Nordlinger est professeur des universités - praticien hospitalier, chirurgien, spécialiste des cancers digestifs. Il a dirigé le pôle digestif de l'Hôpital Ambroise Paré, et le comité cancer de l'APHP (Assistance publique hopitaux de Paris). Ses travaux à l'INSERM ont porté sur la transplantation d'hépatocytes et les modèles de cancers. En cancérologie, les essais cliniques nationaux et internationaux qu'il a organisés et coordonnés, ont permis notamment d'établir le standard international actuel de traitement médico-chirurgical des métastases hépatiques des cancers du colon. Il coordonne avec Cédric Villani le groupe de travail bi-académique « Mathématiques, big data et cancérologie ». Il est membre de l'Académie nationale de médecine.



Cédric VILLANI

Professeur à l'université Claude Bernard Lyon 1, Cédric Villani a été le lauréat de la Médaille Fields en 2010. Ses travaux scientifiques ont ouvert des voies nouvelles dans trois domaines des mathématiques : équation de Boltzmann, transport optimal et géométrie, amortissement de Landau.

Il était directeur de l'Institut Henri Poincaré de l'université Pierre-et-Marie-Curie jusqu'en 2017.

Elu député en 2017 dans la cinquième circonscription de l'Essonne, Cédric Villani est le premier Vice-Président de l'Office parlementaire d'évaluation des choix scientifiques et technologiques (OPECST).

P rogramme

- 14:30** Ouverture du colloque
Sébastien CANDEL, président de l'Académie des sciences, Paris
Claude JAFFIOL, président de l'Académie nationale de médecine, Paris
Pascale COSSART, secrétaire perpétuel de l'Académie des sciences, Paris
Daniel COUTURIER, secrétaire perpétuel de l'Académie nationale de médecine, Paris
- Modérateurs de la séance : **Nicholas AYACHE** et **Bernard NORDLINGER**
- 14:35** Introduction
Cédric VILLANI, **Nicholas AYACHE**, membres de l'Académie des sciences
- 14:45** L'ouverture des données publiques de santé en France
Emmanuel BACRY, directeur de recherche au CNRS à l'université Paris-Dauphine, professeur associé à l'École polytechnique
- 15:15** Modélisation des tumeurs à partir des images radiologiques
Thierry COLIN, professeur à l'Institut polytechnique de Bordeaux
- 15:45** Apprentissage statistique pour la médecine de précision en cancer
Chloé-Agathe AZENCOTT, chargée de recherche, Mines Paris Tech
- 16:15** La recherche clinique en cancérologie à l'ère des nouvelles technologies
Jean-Yves BLAY, directeur général du Centre Léon Bérard
- 16:45** Les messages à emporter
Bernard NORDLINGER, membre de l'Académie nationale de médecine

Résumés et biographies



Emmanuel BACRY

Directeur de recherche au CNRS à l'université Paris-Dauphine, professeur associé à l'École polytechnique

Emmanuel Bacry est directeur de recherche à l'université Paris-Dauphine et professeur associé à l'École polytechnique où il dirige l'Initiative Data Science, dont un des thèmes centraux est le « Big data appliqué à la santé publique et à l'économie de la santé ». Ancien élève de l'École normale supérieure Ulm, docteur en mathématiques et habilité à diriger des recherches, il est actuellement responsable d'un partenariat entre l'École polytechnique et la Caisse nationale d'assurance maladie (CNAMTS) consistant à développer de nouveaux algorithmes d'apprentissage automatique (*machine learning*) sur la base SNIIRAM.

L'ouverture des données publiques de santé en France

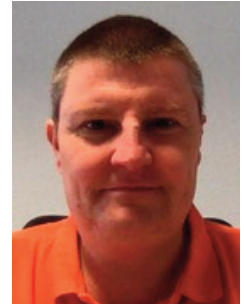
Le SNIIRAM constitue certainement l'une des plus grandes bases de données santé au monde. Au coeur de cet entrepôt de données se trouvent bien sûr la base hospitalière PMSI (Programme de médicalisation des systèmes d'information), mais surtout le DCIR (Consommation inter-régime) qui rassemble l'essentiel des informations du système public de remboursement des soins en France. Une base unique, rassemblant des données sur plus de 65 millions d'individus, plusieurs centaines de Tera-octets, une mine d'or d'informations avec des impacts potentiels forts aussi bien en santé qu'en économie de la santé. Depuis bientôt trois ans, la CNAMTS et l'Ecole polytechnique se sont associés pour déployer de nouvelles pistes d'exploitations du SNIIRAM, pistes issues des dernières avancées tant du point de vue de l'infrastructure que du point de vue des algorithmes mathématiques d'apprentissage statistique. Nous présenterons les premiers résultats obtenus dans le cadre de ce partenariat sur la détection de signaux faibles ou d'anomalies en pharmaco-épidémiologie.



Thierry COLIN

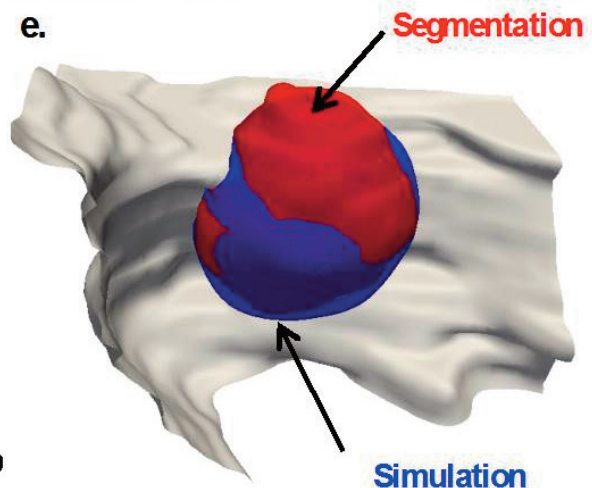
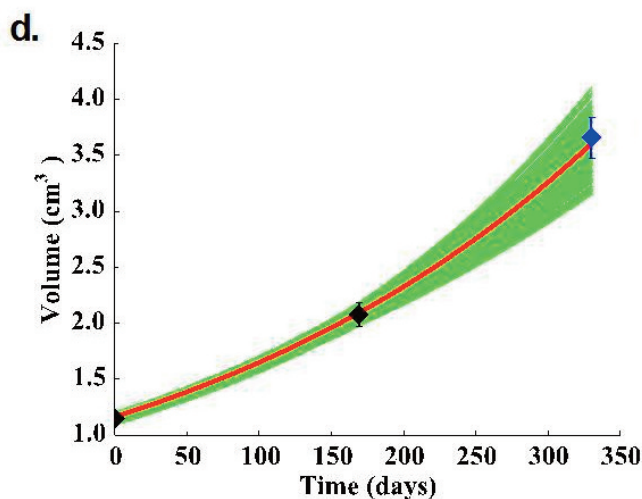
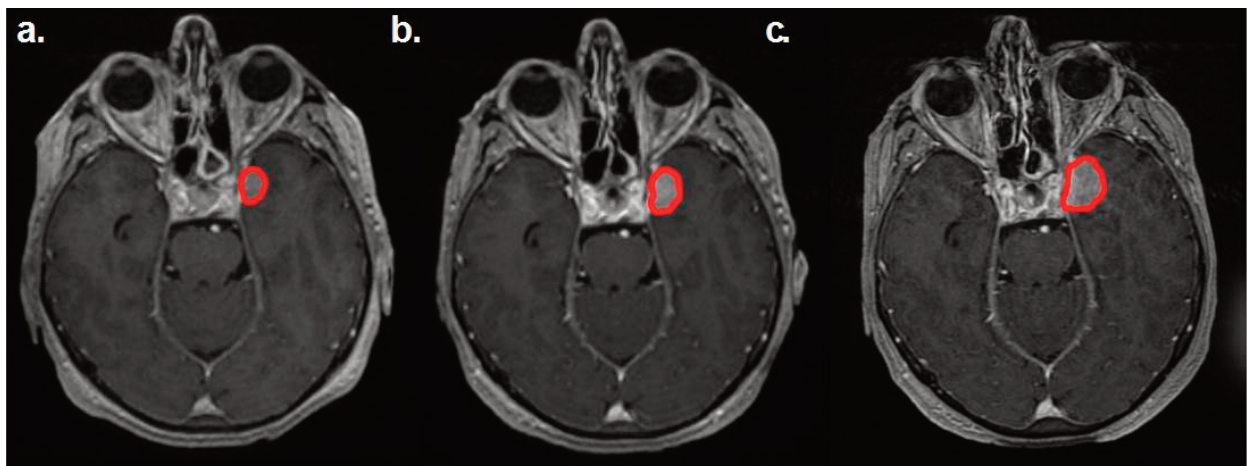
professeur à l'Institut polytechnique de Bordeaux

Ancien élève de l'Ecole normale supérieure de Lyon, Thierry Colin a travaillé au CNRS et à l'université de Bordeaux. Il est actuellement professeur de mathématiques à Bordeaux INP où il enseigne dans plusieurs écoles d'ingénieurs. En début de carrière, il a été conseiller scientifique au CEA et travaillait sur des problèmes d'électromagnétisme et de mécanique des fluides. Depuis 10 ans, ses travaux portent sur la modélisation du cancer et des traitements en collaboration avec des médecins de l'Institut Bergonié et du Centre hospitalier universitaire (CHU) de Bordeaux avec comme but de mettre au point des outils utilisables dans la pratique clinique.



Modélisation de l'évolution des tumeurs à partir d'images radiologiques

L'imagerie joue un rôle central dans la pratique clinique en oncologie que ce soit pour le diagnostic, l'évaluation de l'efficacité du traitement ou bien la surveillance. Pour chaque patient, les cliniciens disposent d'une série d'images. Il est donc extrêmement naturel de tenter de modéliser la dynamique d'évolution de la maladie et des réponses au traitement à partir de ces données d'imagerie. Les simulations et les modèles peuvent potentiellement être utiles dans le cadre d'une surveillance (par exemple quand doit-on opérer un méningiome ?) ou pour anticiper un échappement à un traitement dans le cadre des thérapies ciblées. La démarche de modélisation passe par une collaboration étroite avec les oncologues, chirurgiens et radiologues pour mettre au point des modèles qui contiennent suffisamment d'information biologique et clinique, mais qui soient d'une complexité compatible avec les données d'imagerie disponibles en routine. Dans cette intervention, je montrerai quelques exemples de telles modélisations de la croissance tumorale et je présenterai des résultats obtenus en avec des médecins bordelais sur des métastases pulmonaires, des méningiomes et des tumeurs du rein.





Chloé-Agathe AZENCOTT

Chargée de recherche, Mines Paris Tech

Chloé-Agathe Azencott est diplômée (2005) de l'École nationale supérieure des télécommunications de Bretagne. Après un doctorat portant sur le développement de méthodes d'apprentissage statistique pour la recherche thérapeutique, obtenu en 2010 à l'université de Californie d'Irvine (États-Unis), elle s'est tournée vers l'analyse de données génomiques pendant son séjour postdoctoral aux Instituts Max Planck de Tübingen (Allemagne). Depuis 2013, elle est chargée de recherche au Centre de bioinformatique de MINES ParisTech. Elle enseigne aussi à CentraleSupélec ainsi que pour OpenClassrooms.

Apprentissage statistique pour la médecine de précision en cancer

Les différences génomiques entre patients expliquent en grande partie pourquoi ceux-ci vivent la même maladie différemment. Comprendre ces différences permet de mettre en place une médecine dite personnalisée, dans laquelle le traitement est adapté au génome du patient. Les développements récents des technologies de séquençage à haut débit permettent d'accumuler un volume considérable de données moléculaires pour de grands nombres d'individus. L'analyse statistique de ces données permet d'identifier les régions du génome associées avec la prédisposition, le pronostic ou la réponse thérapeutique. Cet exposé donnera une vue d'ensemble des défis que présentent ces analyses, et des bénéfices que des approches d'apprentissage statistique peuvent apporter.





Jean-Yves BLAY

Directeur général du Centre Léon Bérard

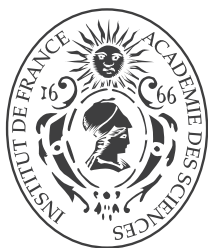
Professeur en oncologie à l'université Claude Bernard Lyon 1, chercheur au Centre de Recherche en Cancérologie de Lyon (CRCL), Jean-Yves Blay est directeur général du Centre Léon Bérard, centre de lutte contre le cancer de Lyon et de Rhône-Alpes depuis 2014.

Par ailleurs, il anime l'*European network for Rare adult solid Cancer* (Euracan), est le coordinateur du réseau de références cliniques spécialisées en sarcomes des tissus mous et des viscères (Netsarc) et préside le Groupe Sarcome Français.

La recherche clinique en cancérologie à l'ère des nouvelles technologies

La médecine moléculaire du cancer s'appuie sur l'identification d'anomalies génomiques de l'ADN des cellules tumorales pour guider le traitement des patients, en choisissant des inhibiteurs agissant sur les protéines codées par les gènes mutés. On assiste actuellement à une évolution d'une vitesse inédite de nos connaissances sur les anomalies génomiques des cellules cancéreuses au sein des programmes nationaux ou internationaux comme ICGC ou TCGA, permettant de déboucher sur de nouvelles classifications nosologiques mais démontrant aussi l'extrême complexité et variabilité clonale des cellules cancéreuses chez l'homme. Il s'agit désormais d'utiliser ces données nouvelles. Ceci requiert la constitution de plateformes de diagnostic explorant progressivement avec une plus grande profondeur les anomalies moléculaires de chaque patient individuel, et l'organisation de réunions de concertation pluridisciplinaires moléculaire permettant d'intégrer ces données au contexte clinique et global du patient, avec des bio-informaticiens. La recherche clinique en cancérologie doit faire évoluer ses paradigmes pour faire face à cet afflux rapide de connaissances.





INSTITUT DE FRANCE
Académie des sciences



Composition du groupe de travail bi-académique "Mathématiques, big data et cancérologie" :

Nicholas AYACHE, Emmanuel BACRY, Martine BENAMAR, Mehdi BENCHOUFI, Jean Yves BLAY, Jeanne BOSSI-MALAFOSSE, Claude Henri CHOUARD, Thierry COLIN, François DUBOIS, Olivier de FRESNOYE, Florence HUBERT, Claude HURIET, Jean Patrick LAJONCHERE, Jean-Roger LE GALL, Bertrand LUKACS, Cécile MONTEIL, Patrick NETTER, **Bernard NORDLINGER (Coordonnateur)**, Patrice QUENEAU, Philippe RAVAUD, Jacques ROUESSE, Guy VALLANCIEN, Alain-Jacques VALLERON, **Cédric VILLANI (Coordonnateur)**, Richard VILLET, Jean-Philippe VERT.

