

# Les logiques politiques de l'ouverture des données de santé en France



François Briatte<sup>1</sup> et Samuel Goëta<sup>2</sup>

Doctorants

L'open data semble pleine de promesses pour le système de santé : sécurité sanitaire, transparence des dépenses de santé et des tarifs médicaux... Armée de ces arguments, une coalition hétéroclite d'acteurs fait aujourd'hui pression sur les pouvoirs publics pour obtenir l'ouverture des données de santé. Mais quelles données exigent-ils ? Pour quoi faire ? Et surtout, quel peut être l'impact de l'open data sur la régulation du système de santé français ?

En France, où le scandale du 'Mediator' a incité les pouvoirs publics à réexaminer les moyens de la surveillance pharmaco-épidémiologique, la "libération des données" a récemment pris la forme d'un débat sur l'accès aux données de santé, en particulier aux fichiers anonymisés des caisses d'assurance-maladie, mesure étudiée dans un rapport récent de l'Inspection générale des affaires sociales<sup>3</sup>.

Ces données, qui proviennent principalement du système d'information inter-régimes de l'assurance-maladie (SNIIRAM), fournissent une information extrêmement détaillée sur le remboursement des prescriptions de la médecine de ville, ce qui a pour effet de rendre mesurable la consommation médicale à l'échelle d'un territoire donné, et de rendre ainsi visible l'activité (et le coût) des professionnels de santé.

Les revendications d'une multitude d'acteurs exigeant un accès facilité à ces données font allusion directement à la notion "d'open data"<sup>4</sup>. Cette dernière fait référence à une injonction de la société civile et à une obligation légale émergente qui vise à établir la mise à disposition de données publiques, anonymisées et non personnelles, de manière proactive par les administrations. L'open data facilite, techniquement et légalement<sup>5</sup>, la réutilisation des données telles que les administrations les détiennent par quiconque, sans exiger de connaître l'identité ni les objectifs de l'utilisateur<sup>6</sup>. Ce dernier point est crucial dans le domaine de la santé où l'utilisateur de données doit souvent décliner patte blanche avant d'accéder aux données même anonymisées. Dans une démarche mêlant transparence et innovation, les politiques d'open data visent à développer une meilleure connaissance de la société et de l'action publique par les multiples médiations créées par la société civile, journalistes de données et entreprises en tête.

1. Doctorant à l'Institut d'Études Politiques de Grenoble et à l'European School of Political Sciences, Lille ; [f.briatte@ed.ac.uk](mailto:f.briatte@ed.ac.uk).
2. Doctorant à Telecom ParisTech, département de sciences économiques et sociales, Paris ; [samuel.goeta@telecom-paristech.fr](mailto:samuel.goeta@telecom-paristech.fr).
3. Pierre-Louis Bras, Rapport sur la gouvernance et l'utilisation des données de santé, Paris, Ministère des affaires sociales et de la santé, 2013.
4. Laetitia Clavreul, "Bataille autour de l'accès aux données de santé", Le Monde, 31 janvier 2013.
5. "Ten Principles for Opening Up Government Information", Sunlight Foundation, Sebastopol (CA), 11 août 2010.
6. Cet élément est mis en exergue dans "l'Open Definition" produite par l'Open Knowledge Foundation : <http://opendefinition.org/>.

Or, bien que ce mouvement soit désormais bien ancré suite à la signature d'une charte pour l'open data par les dirigeants du G8 de 2013, l'accès aux données de santé obéit à toute une série de restrictions légales et techniques qui limitent cette réutilisation<sup>7</sup>. De telles réticences ont été exploitées par l'Initiative Transparence Santé, un regroupement d'acteurs hétéroclites dont nous aborderons la composition et les revendications, qui s'interroge dans sa pétition sur les "choses à cacher" que protégeraient les pouvoirs publics ou une éventuelle crainte de voir des dysfonctionnements mis en lumière. Désignant clairement un sens de l'histoire, leur manifeste publié en mars 2013 regrette que la France soit à "contre-courant" à "l'heure de l'open data"<sup>8</sup>.

Néanmoins, au-delà de l'accès aux données, la controverse en cours sur les données de santé interroge plus largement la configuration du système de santé français. Historiquement, le terme open data est associé à celui d'open government qui, par un accès plus large à l'information, vise à étendre la participation de la société civile dans l'élaboration des politiques publiques<sup>9</sup>. L'ouverture des données peut alors se comprendre comme une opportunité, pour les acteurs d'un secteur de l'action gouvernementale, de renégocier avec l'État l'étendue de sa régulation, aussi bien en termes de participation de la société civile aux politiques publiques qu'en termes d'accès à un marché comme celui de la santé par des sociétés privées.

## La mobilisation de l'Initiative Transparence Santé

L'initiative Transparence Santé (ITS) a été lancée le 25 janvier 2013 et se présente comme un regroupement d'acteurs sans rattachement à une structure formelle. Parmi ses premiers initiateurs, on trouve des personnalités bien insérées dans le milieu de la santé comme Thomas Laurenceau, Rédacteur en chef de 60 millions de consommateurs, Alain Bazot, Président de l'UFC-Que-Choisir ou Christian Saout, ancien Président du Collectif interassociatif sur la Santé (Ciss), un regroupement d'associations de patients. On y trouve aussi des dirigeants d'entreprises de taille moyenne dont le cœur de métier dépend de la diffusion des données de santé.

Parmi celles-ci, Celtipharm, une entreprise qui se décrit comme "spécialiste de la conception et de la vente de programmes marketing-ventes pour l'industrie pharmaceutique" est un de ses membres les plus actifs. La société a exercé en 2013 une intense activité de lobbying auprès du ministère de la santé car l'accès aux données de la santé est son fonds de commerce. En plus de réclamer l'accès aux données publiques, Celtipharm tente de faire valoir l'autorisation accordée par la CNIL en 2011 de collecter des données anonymisées sur le contenu des ordonnances. Il s'agissait de réaliser des "études épidémiologiques à partir de données issues des feuilles de soins électroniques anonymisées". L'entreprise se heurte aux caisses d'assurance maladie et aux mutuelles, qui refusent de mettre à disposition les clefs de déchiffrement indispensables à l'analyse des flux de données chiffrées<sup>10</sup>. La mobilisation de l'entreprise auprès des pouvoirs publics pour exiger de décrypter les télétransmissions des ordonnances a valu à la société d'être nominée aux "Big Brother Awards" de l'ONG Privacy France<sup>11</sup>.

Une autre entreprise, Fourmi Santé, mérite d'être citée parmi les acteurs à l'initiative du mouvement. En septembre 2012, l'entreprise a fait l'objet d'une intense couverture médiatique<sup>12</sup>. Elle réutilisait les données mises à disposition par la Caisse nationale de l'assurance maladie des travailleurs salariés (CNAMTS) sur le site ameli.fr. Le site permettait aux utilisateurs de sélectionner les médecins, notamment en fonction des prix moyens des consultations disponibles sur la base de données. La CNAMTS a bloqué l'extraction automatique des données

7. Une récente enquête sur les dépassements d'honoraires à Paris illustre l'important travail à accomplir afin d'obtenir une vue d'ensemble des tarifs des consultations médicales : Jean-Baptiste Chastand, Laetitia Clavreul et Alexandre Léchenet, "Tarifs médicaux : enquête sur les dépassements d'honoraires", Le Monde, 10 avril 2012.

8. Manifeste du collectif Open Data Santé, mars 2013.

9. Harlan Yu et David G. Robinson, "The New Ambiguity of 'Open Government'", UCLA Law Review Discourse, 178, 2012.

10. Guillaume Champot, "CeltiPharm veut déchiffrer les ordonnances transmises à la Sécu", Numerama, 7 mars 2013.

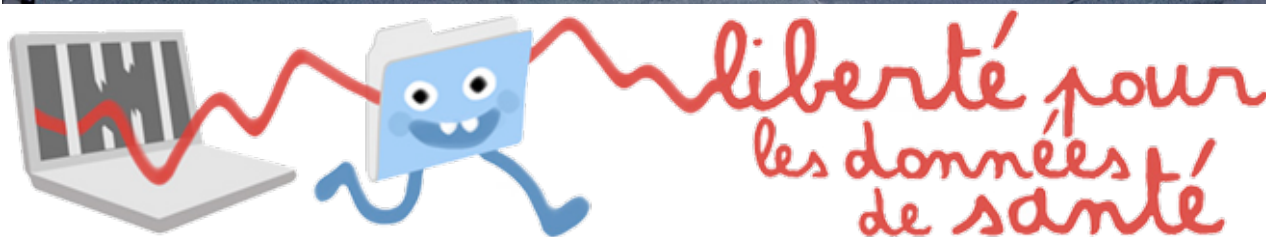
11. "Celtipharm", Big Brother Awards 2013.

12. AFP, "Honoraires médicaux: la "Sécu" interdit à un site de comparer les tarifs", 5 septembre 2012.

qu'effectuait l'entreprise et a émis une mise en demeure exigeant le retrait des liens vers le site [ameli.fr](http://ameli.fr)<sup>13</sup>. Fourmi Santé a alors répliqué par une intense campagne médiatique puis a joué un rôle moteur dans la création de l'Initiative Transparence Santé.

L'Initiative Transparence Santé est vite parvenue à imposer son message dans l'espace public par une stratégie d'influence efficace. Elle s'appuie sur une intense activité éditoriale à travers la publication du blog [opendatasante.com](http://opendatasante.com) et de tribunes dans le journal Le Monde et le site d'information libéral Contrepoints. Le message de l'Initiative est aussi portée par la société civile avec une pétition signée par près de 2500 personnes. Une rapide analyse sémantique des professions déclarées par les signataires sur le site [opendatasante.com](http://opendatasante.com) montre une très forte proportion de professionnels de santé et environ 30% de retraités. Mobiliser ce dernier public sur une problématique technique comme celle de l'open data témoigne d'une organisation structurée des initiateurs de la démarche et de relais bien établis dans les associations de patients et de consommateurs susceptibles de porter le message auprès de ce public.

La pétition exige une diffusion de l'ensemble des données de santé dans une définition très large de la "manne d'informations relatives à l'utilisation et au fonctionnement de l'offre de soins". Les données sont représentées visuellement comme des détenus à libérer, une image de la Prison de la Santé à Paris illustre la pétition en arrière-plan. La pétition donne quelques exemples de cette "manne" à libérer : "tarifs des professionnels, des produits de santé, des hôpitaux, informations relatives à la qualité et à la consommation des soins, rapports d'activité des caisses primaires d'assurance maladie, efficacité des médicaments, qualité de la prescription"<sup>14</sup>. En somme, le manifeste relaie une exigence d'ouverture de l'ensemble des données publiques relatives au domaine de la santé, dans la lignée du mouvement plus large d'ouverture des données publiques.



13. "La Sécurité sociale censure Fourmi Santé, un site comparateur de tarifs médicaux", Slate.fr, 2013.

14. Manifeste du collectif Open Data Santé, op. cit.



## L'objet du litige : la base SNIIRAM

Passé le manifeste qui fait office de vitrine de l'initiative, les tribunes publiées sur [opendatasante.com](http://opendatasante.com) font part d'une demande bien plus ciblée qu'une simple volonté d'ouvrir les données de santé dans leur ensemble. La mission Etalab sous l'autorité du Premier Ministre, en charge de l'ouverture des données publiques de l'Etat, a entrepris un ambitieux recensement des bases de données disponibles en matière de santé produites par l'Etat mais aussi des fondations, des entreprises et même le gouvernement américain<sup>15</sup> ! Près de 280 données produites par 58 organismes sont recensées dans le tableur publié par Etalab. Dans son travail de recensement, Etalab a évalué l'ouverture des données de santé qu'elle a pu identifier selon quatre critères principaux qui correspondent aux exigences les plus communes des militants de l'open data : la liberté d'accès aux données par le grand public plutôt qu'une catégorie de personnes, la gratuité d'accès, la mise à disposition des données dans un format exploitable et enfin la possibilité de réutiliser les données sans restriction. Tous ces critères ont été évalués à la fois pour les données au niveau granulaire défini comme le "niveau le plus fin qu'il est possible d'obtenir en fonction de l'origine de la donnée et du système de collecte" ou agrégé soit le niveau obtenu "en regroupant des données granulaires selon une ou plusieurs caractéristique(s) commune(s)". Le recensement entrepris par Etalab est toujours en cours, de nombreux indicateurs ne sont pas encore renseignés.

Critère	Données agrégées	Données granulaires
Liberté d'accès	Accessible à tous : 95 En accès restreint : 2 Non précisé : 189	En accès restreint : 154 Accessible à tous : 47 En accès fermé : 19 Non précisé : 66
Coût d'accès	Gratuit : 97 Non précisé : 189	Gratuit : 138 Payant : 40 Payant ou gratuit : 10 Non précisé : 98
Format d'accès	Non exploitable : 73 Exploitable : 22 Non précisé : 191	Exploitable : 163 Non exploitable : 25 Non précisé : 98
Condition de réutilisation	Explicité avec restriction : 57 Explicité sans restriction : 22 Non explicité : 9 Non précisé : 198	Explicité avec restriction : 107 Explicité sans restriction : 10 Non explicité : 14 Non précisé : 155

Figure 1. Synthèse de la cartographie des données de santé réalisée à partir du fichier mis à disposition par Etalab (avril 2014)

Comme l'illustre le tableau ci-dessus, les restrictions d'accès se concentrent principalement sur les données granulaires. Ce sont en effet celles-ci qui concentrent les difficultés d'anonymisation, essentielles pour éviter la ré-identification des patients. Ce sont aussi ces données difficilement accessibles qui intéressent le plus les entreprises, les militants et les chercheurs à l'origine de l'Initiative Transparence Santé.

Si l'on cumule les quatre critères d'ouverture des données au niveau le plus proche des revendications des militants de l'open data, on constate que 13 données agrégées et 9 données granulaires sont d'ores et déjà disponibles. La plateforme [data.gouv.fr](http://data.gouv.fr) propose ainsi des

15. Etalab, "Cartographie des bases de données publiques en santé", avril 2014.

données comme les honoraires moyens des professionnels de santé, les indicateurs relatifs aux infections nosocomiales, ou encore la base complète [des médicaments en vente en France](#). Ces données déjà disponibles, l'Initiative Transparence Santé ne les évoque pas tout comme elle n'a pas réagi à la cartographie des données de santé en dépit du fait que cette étude serait précieuse à ses revendications.

L'Initiative concentre son action sur un acteur en particulier, la CNAMTS. Depuis mars 2013, peu après la publication de sa tribune fondatrice, l'Initiative a précisé ses exigences dans un billet sur [opendatasante.com](#) qui dénonce la mauvaise exploitation de la base SNIIRAM. L'article s'appuie sur un rapport de l'Institut des Données de Santé qui décrit l'absence d'accès direct à la base pour des institutions comme la Haute Autorité de Santé ou l'Agence Nationale de Sûreté du Médicament. Le plaidoyer est clair dans ses ambitions : "l'accès à ces données devrait être élargi à l'ensemble de la société civile et non restreint à quelques privilégiés sous la houlette d'un ministère de tutelle décidant qui a le droit d'obtenir de l'information et qui en sera privé."<sup>16</sup> Que contient cette base tant convoitée ? Selon le rapport de l'inspecteur général des affaires sociales Pierre-Louis BRAS sur la gouvernance et la diffusion des données de santé adressé à la ministre de la Santé Marisol Touraine en septembre 2013, le SNIIRAM tire initialement ses données des feuilles de soins, environ 1,2 milliard d'enregistrements par an qui sont anonymisés selon une procédure stricte. Mise en place par une loi de 1998 dans le Code de Sécurité Sociale et en fonctionnement depuis 2003, la base SNIIRAM se compose de données sur les bénéficiaires, les professionnels et les structures de soin, les demandes de prise en charge et leur remboursement et enfin de données sur les hospitalisations. Les données sont consultables pour une période de 3 ans et l'année en cours mais sont conservées par la CNAMTS pendant 10 ans, en cas de crise nécessitant une exploration des données dans des périodes antérieures.

En juillet 2013, l'Initiative Transparence Santé change de tactique pour enfin mettre la main sur les données de la base SNIIRAM. Elle adresse une requête à la CNAMTS pour obtenir des données sur la consommation de Mediator s'appuyant sur la loi de 1978 qui donne un droit d'accès aux informations publiques : "En juillet, notre collectif a officiellement saisi la CNAMTS d'une demande relative à la consommation de Mediator. Quelles quantités ont été consommées ? Dans quelle mesure les prescriptions étaient médicalement justifiées ? Et surtout, combien la collectivité a-t-elle dépensé afin de rembourser l'empoisonnement de centaines de patients ?"<sup>17</sup>. La CADA (Commission d'Accès aux Documents Administratifs) créée par la loi de 1978 a émis un avis favorable à ce que les données soient communiquées à l'Initiative Transparence Santé : "la communication de ces informations au collectif sous la forme demandée par celui-ci, n'est pas de nature à porter atteinte au secret médical ou au secret en matière commerciale et industrielle"<sup>18</sup>. Une nouvelle requête a donc été formulée à la suite de cet avis et la CNAMTS a adressé sa réponse dans une lettre comportant les données demandées, notamment l'effectif des assurés par département de 1999 à 2009 à qui du Mediator a été prescrit, les spécialités des médecins à l'origine des prescriptions et l'évolution des montants remboursés.

Cette opération a permis à l'Initiative de démontrer les possibilités de l'ouverture des données de santé. Avec les données du Mediator, l'Initiative déploie son argument le plus fort : que la transparence et la plus grande redevabilité (accountability) du système de santé rendues possibles par l'open data permettra d'éviter les prochains scandales sanitaires. L'initiative décrit des scénarios d'usage des données de santé dans lesquels quiconque peut "suivre l'utilisation d'un médicament [...] et mesurer si les recommandations d'usage, émises par les agences sanitaires, n'ont pas été ignorées ou détournées"<sup>19</sup>. Mais, dans les arguments de l'Initiative, il s'agit au delà de reconfigurer la régulation du système de santé en encourageant « le développement de la culture du débat contradictoire et de la contre expertise » reprenant l'expression de l'Institut des Données de Santé. Le comité d'experts de cette institution en

16. ITS, "Accès aux données de santé : l'opposition se précise", 20 mars 2013.

17. ITS, "L'open data, la ministre de la Santé n'en veut pas", Le Monde, 2013.

18. AFP, "Avis favorable pour plus de transparence des données de santé", 30 décembre 2013.

19. ITS, "Débattre d'urgence de l'accès aux données de santé", Le Monde, 30 janvier 2013.

charge du partage et de la diffusion des données entre organismes de santé s'est en effet récemment prononcé en faveur d'une plus large diffusion des données de santé, notamment en conditionnant l'accès non pas à un statut institutionnel mais à la finalité de l'étude conduite avec les données<sup>20</sup>.

Ainsi, l'ouverture des données de santé amène plus largement à réfléchir à une reconfiguration du système de santé français. En ouvrant les données de la base SNIIRAM, il s'agit de donner la possibilité à un plus large panel d'acteurs - sociétés privées, chercheurs, citoyens experts - de participer au contrôle du système de santé. Une ouverture qui remet en cause les rapports d'acteurs établis progressivement dans les institutions de santé.

## La configuration du système de santé français<sup>21</sup>

Le système français de protection sociale s'est construit sur un rapport de force constitutif : bien que d'inspiration universaliste, sa généralisation n'a pas abouti à la création d'un régime unique de sécurité sociale, ou d'un "service national de santé" à l'image du système de santé britannique. Au contraire, la gestion des prestations sociales y est confiée aux représentants de différentes branches du salariat, réunis au sein de caisses autonomes de l'État et régies par les principes de la gestion paritaire<sup>22</sup>.

Dans le secteur de la santé, cette organisation signifie que la régulation de l'assurance-maladie repose primordialement sur un accord entre deux acteurs non étatiques, représentant l'offre médicale et la demande de soins : les professionnels de santé, dont la représentation syndicale en France est très fragmentée, et les dirigeants des caisses d'assurance-maladie, au sein desquelles la CNAMTS, qui assure quatre personnes sur cinq en France et qui finance actuellement les trois quarts des dépenses de santé, est amenée à jouer un rôle central.

Cette organisation explique pourquoi la "libération" des données de santé en France revient en partie à les ouvrir... aux services de l'État : jusqu'à récemment, les bases de données de la Sécurité sociale n'étaient directement accessibles qu'aux agents de l'assurance-maladie. Cet accès, désormais étendu aux services ministériels, permet à l'État de faire intervenir sa propre expertise des dépenses de santé, sans avoir recours à celles de personnels indépendants.

Vu sous cet angle, "libérer les données" de l'assurance-maladie revient ainsi à supprimer une prérogative des organismes chargés, à la création du système de santé, d'autonomiser la régulation des soins des services de l'État. Cette dynamique d'étatisation graduelle prévaut dans les réformes du système de soins français sur les vingt dernières années, se retrouve également dans d'autres branches de la sécurité sociale<sup>23</sup>, et a plusieurs équivalents fonctionnels dans d'autres systèmes de santé.

Ces changements d'attribution obéissent à deux principaux impératifs pour les pouvoirs publics : réduire la pression fiscale exercée par les branches maladies des États-providence, et prévenir les scandales sanitaires causés par les crises de santé publique<sup>24</sup>. Il pourrait éventuellement en servir d'autres, comme la réduction des inégalités sociales et territoriales de santé, qui ne dispose toutefois ni du sentiment d'inévitabilité associé à la régulation des dépenses de soins, ni du sentiment d'urgence concomitant aux crises sanitaires.

20. ITS, "Le comité d'experts de l'Institut des données de santé favorable à un très large accès", 4 mai 2013.

21. Cette section a fait l'objet d'une présentation à l'Apéro Science & Web n°33 en février 2014. Merci aux participants pour leurs questions et leurs réactions.

22. Dominique Damamme et Bruno Jobert, "Les paritarismes contre la démocratie sociale", Pouvoirs, n°94, 2000, p. 87-102.

23. Jean-Claude Laborier et Bruno Théret, Le système français de protection sociale, Paris, La Découverte, 2009.

24. Didier Tabuteau, Démocratie sanitaire, Paris, Fayard, 2013, p. 181.

## Les implications de l'ouverture des données de santé

La "libération" des données de santé a ainsi plusieurs interprétations possibles, qui vont du simple travestissement d'intérêts industriels à l'expression de revendications citoyennes portées par des représentants d'associations de consommateurs. Les tensions relatives au système de soins français expliquent en partie cette situation, dans la mesure où la lisibilité des données de l'assurance-maladie est la résultante d'un rapport de force institutionnel organisé autour de la régulation des soins.

Pour ces raisons, la limitation de l'accès aux données joue, dans le domaine de la santé, le rôle de barrière à l'entrée d'un champ d'action stratégique, dédié au contrôle des dépenses de santé, et où s'élabore la mesure économique de la maladie et des soins. L'élite décisionnelle de ce champ est largement composée de hauts fonctionnaires qui ont intérêt à l'ouverture de ces données, mais prioritairement voire exclusivement pour équiper leurs propres services.

Que signifierait, à terme, l'accès de tous aux données de santé en France ? L'effet premier d'un tel changement serait de faire baisser drastiquement les coûts d'analyse de l'activité clinique pour les acteurs déjà impliqués dans cette entreprise, au premier rang desquels les acteurs industriels de l'assurance et des produits de santé. L'effet espéré d'une modernisation de la surveillance épidémiologique en France est quant à lui à mettre en perspective avec la réaction des professionnels de santé à l'examen "ouvert" de leur activité par quiconque, plutôt qu'au moyen des dispositifs volontaires existants mis en place au sein des Agences Régionales de Santé<sup>25</sup>.

Comprise dans le contexte plus général de la réforme des États-providence européens, la libéralisation de l'accès aux données de santé faciliterait l'évaluation économique de la consommation médicale pour les acteurs publics, tout en rendant les acteurs privés plus à même de participer à la compétition autour de l'offre dans ce domaine. Cette dualisation caractérise de nombreuses réformes de régimes de protection sociale, recommandées par des institutions comme la Banque mondiale et visant à établir des systèmes mixtes public-privé de couverture assurantielle<sup>26</sup>.

L'ouverture des données repose toutefois sur une prémisses invitant à définir un enjeu supplémentaire : l'accès généralisé aux données permettrait à de nouveaux entrants d'investir le champ d'action stratégique de la santé en remettant en question la qualité des décisions qui y sont prises, chiffres et données à l'appui. Dans cet ordre d'idée, il est peu surprenant d'apercevoir des sociétés de services aux industries assurantielles et pharmaceutiques au premier rang de ceux appelant à l'émergence d'une contre-expertise en matière de santé.

Le projet d'une contre-expertise citoyenne, "statactiviste"<sup>27</sup>, dans le secteur de la santé, ferait quant à lui écho au projet de la démocratie sociale, mais aussi à celui de la démocratie sanitaire, dans la mesure où il viserait à permettre aux intéressés eux-mêmes d'infléchir responsablement sur le déroulement et l'organisation des soins, en plaçant entre leurs mains les ressources nécessaires pour en décider collectivement. Ce projet n'a néanmoins que très peu de prise sur la réalité pratique des demandes auxquelles les pouvoirs publics sont aujourd'hui confrontés.

25. Ces dispositifs désignent les "contrats" passés entre les pouvoirs publics et les médecins libéraux, et qui permettent notamment à l'État d'inciter les praticiens à prescrire des médicaments génériques et à se rendre dans des zones géographiques de faible démographie médicale.

26. Bruno Palier, *Gouverner la Sécurité sociale*, Paris, Presses Universitaires de France, coll. "Quadrige", 2005.

27. Isabelle Bruno, Julien Prévieux et Emmanuel Didier (dir.), *Statactivisme, comment lutter avec des nombres*, Paris, La Découverte, coll. "Zones", à paraître.