

ARTICLE

**Numéro spécial : patients partenaires, patients experts.
Quels dispositifs de participation en santé ?**

**Questions médicales controversées, déclarations
de consensus et participation du public : le cas
des conférences de consensus du National
Institute of Health**

Stéphanie Debray

Archives Henri-Poincaré — Philosophie et recherches sur les sciences et les technologies (AHP-PreST),
Université de Lorraine, Nancy, France
Auteure-ressource. Courriel : stephanie.debray@univ-lorraine.fr

Résumé

Les conférences de consensus du *National Institute of Health* (NIH), aujourd'hui abandonnées, ont servi de modèle pour établir des programmes de développement de consensus similaires dans d'autres pays. L'intérêt épistémique de ce dispositif fut notamment discuté. Cet article vise à examiner deux positions philosophiques opposées au moyen d'une synthèse des arguments adressés à Miriam Solomon par Laszlo Kosolosky et Jeroen Van Bouwel. Une position philosophique intermédiaire laisse ensuite entrevoir l'intérêt méthodologique d'une analyse rétrospective des conférences de consensus du NIH pour évaluer l'impact d'un forum public sur les recommandations *in fine* transmises aux professionnels de santé et citoyens.

Abstract

The now retired NIH Consensus Development Program has been used as a model for similar programs in other countries. However, the epistemic value of this kind of program has been disputed. This article provides an overview of the arguments levelled at Miriam Solomon by Laszlo Kosolosky and Jeroen Van Bouwel, who provide an opposing philosophical position on this issue. Here, I argue for a middle ground position that highlights the methodological interest of a retrospective analysis of the NIH consensus conferences as a way to assess the impact of a public forum on guidelines ultimately relayed to health-care professionals and citizens.

Mots-clés : science et public ; consensus scientifiques ; conférences de consensus ; pratique clinique ; éthique médicale

1. Introduction

Il est déjà admis, en particulier dans le domaine des études sur les sciences et les technologies (Science and Technology Studies — STS), que de nombreuses raisons, principalement éthiques et démocratiques, peuvent inciter les scientifiques à prendre en considération les discours non académiques au sein des décisions scientifiques, ou encore à inclure le public dans la pratique scientifique¹. Cela se traduit depuis les années soixante-dix par l'apparition, puis la multiplication, des sciences dites aujourd'hui « citoyennes », « participatives », ou « collaboratives »². La participation du public à l'activité scientifique peut prendre différentes formes, et ce à différents moments : lorsqu'il s'agit de choisir un sujet d'étude, de choisir la méthodologie appropriée, de récolter des données, ou encore lorsqu'il s'agit de communiquer des résultats. Dans le domaine de la santé, le public — pouvant être ici composé de patients, de patients experts, de représentants de patients, d'usagers, de citoyens — peut parfois participer à différents niveaux : en ayant un rôle de médiateur, en participant à des groupes de travail et à des commissions, en étant impliqué dans l'administration des établissements de santé ou dans la formation des professionnels de santé, en participant à des débats publics ou des conférences de consensus. Néanmoins, lorsque la science est perçue comme une activité devant être exempte de valeurs non épistémiques au moment du raisonnement scientifique et autonome vis-à-vis des autres sphères de la société, la participation des citoyens à l'activité scientifique est rarement perçue comme une nécessité et les valeurs non épistémiques peuvent à l'inverse nuire à l'intégrité et à l'objectivité de la recherche³. En outre, la possibilité que des échanges avec les citoyens puissent, sur le plan épistémique, permettre de résoudre des sujets controversés de façon rationnelle n'est pas assurée. La participation du public, en particulier dans le domaine médical, est alors souvent focalisée sur les éventuels bénéfices pour le patient, son bien-être, sa prise en charge, et pour l'amélioration de la pratique clinique. Mais qu'en est-il des bénéfices que cela peut potentiellement apporter aux scientifiques et à la recherche ? La participation du public permet-elle d'améliorer les résultats scientifiques et les décisions prises par des scientifiques ou chercheurs sur des questions médicales ? En partant du constat selon lequel l'implication du public est souvent abordée du point de vue éthique et démocratique, mais peu étudiée d'un point de vue épistémique, le présent article vise à poser la question suivante : comment savoir si la participation du public a un impact favorable sur des questions médicales et de santé usuellement propres aux scientifiques et aux chercheurs (innocuité, efficacité thérapeutique, épidémiologie, directions de la recherche, recommandations cliniques sur les traitements et sur la prise en charge, interprétation des preuves) ? Les conférences de consensus sont l'un des endroits où scientifiques et public sont réunis pour débattre et

¹ Cet article a pour origine une communication présentée lors de journées d'études organisées par Claire Crignon, Renaud Debailly et Antoine Guillain sur le sujet suivant : « Patients-experts, patients-partenaires : quels dispositifs de participation en santé ? » (Debray, 2021).

² Ces expressions allant souvent de pair peuvent néanmoins être distinguées en fonction des modalités de participation du public et des types de public concernés.

³ L'ouvrage *Science, Policy, and the Value-Free Ideal* de Heather Douglas (2009), et de manière générale la littérature au sujet de l'idéal d'une science exempte de jugements de valeurs, permettent d'approfondir cet aspect.

prendre des décisions sur des questions communes. Un second constat est alors le suivant : les conférences de consensus de l'Institut américain de la santé⁴ (National Institute of Health — NIH) font partie des premières conférences de consensus du domaine médical à inclure le public au dispositif de participation, tant sur des questions « techniques » qu'éthiques ou politiques. Ces conférences de consensus pourraient en ce sens être un objet d'étude adapté pour évaluer l'impact d'une participation du public sur des questions médicales et de santé habituellement réservées aux scientifiques et aux chercheurs. Ce programme de développement de consensus (Consensus Development Program — CDP), et donc les conférences de consensus en question, a été supprimé en 2013 par le Bureau de la prévention des maladies (Office of Disease Prevention — ODP), mais il a servi et sert encore de modèle pour développer des programmes similaires dans d'autres pays. Focaliser l'attention sur le CDP du NIH ne permet pas d'envisager la question de la participation du public de manière générale ni d'apporter des réponses qui concerneraient « la participation du public » dans son ensemble, pas plus qu'elle ne permet d'obtenir des résultats généralisables. Néanmoins, cela pourrait, vis-à-vis de la question posée, permettre d'en apprendre davantage sur l'impact positif, négatif ou négligeable que peut avoir un forum public sur des questions discutées généralement entre scientifiques et entre chercheurs, ainsi qu'apporter éventuellement des informations (puisque le CDP a servi de modèle) sur des dispositifs similaires de participation du public en santé. Le présent article vise alors plus précisément à envisager la question de la participation du public, en particulier dans le domaine de la santé, en se focalisant sur son intérêt (ou son absence d'intérêt) épistémique, et en prenant comme objet d'étude les conférences de consensus du NIH. La question est donc de savoir si ce CDP, *en tant que dispositif qui incluait le public*, possédait (ou non) un intérêt du point de vue épistémique. Cela demande en réalité de répondre au moins à deux questions. Il s'agit : i) de déterminer, d'une part, si le CDP en tant que dispositif qui incluait divers types d'acteurs possédait réellement⁵ un intérêt épistémique ; ii) dans l'affirmative, de déterminer si cela était dû d'une quelconque manière à la participation du public.

Le premier point fut déjà discuté en philosophie des sciences. Les travaux de Miriam Solomon, en particulier, semblent apporter une réponse négative à ce sujet, notamment dans le chapitre « The Social Epistemology of NIH Consensus Conferences » (Solomon, 2007), puis dans « Group Judgment and the Medical Consensus Conference » (Solomon, 2011)⁶. Cependant, dans l'article « The Intended Window of Epistemic Opportunity: A Comment on Miriam Solomon » (2012), Laszlo Kosolovsky remet en cause les arguments de l'auteure, et dans « Explicating Ways of Consensus-Making in Science and Society: Distinguishing the Academic, the Interface and the Meta-Consensus », Kosolovsky et Jeroen Van Bouwel défendent, contrairement selon eux à la position de Solomon, que ces

⁴ Le pluriel est également utilisé aujourd'hui du fait de la création de plusieurs centres et instituts ; par souci de cohérence avec les articles cités dans le présent texte, la forme du singulier sera plutôt conservée.

⁵ C'était l'un des objectifs visés dès les années soixante-dix, mais cela ne signifie pas nécessairement que cet objectif fut atteint.

⁶ Voir aussi l'ouvrage *Making Medical Knowledge* (Solomon, 2015).

conférences de consensus avaient un intérêt épistémique, et ne servaient pas seulement à disséminer les informations scientifiques auprès du public. Je propose alors, pour répondre au premier point, une synthèse des arguments introduits par ces auteur(e)s (sections 2, 3, et 4), afin de déterminer si ces conférences de consensus du NIH possédaient en définitive un intérêt épistémique.

Le second point demande de s'interroger, à la lumière de cette synthèse, sur ce que ces conférences de consensus — auxquelles pouvaient participer des non-spécialistes et des citoyens — nous apprennent au sujet de la participation du public sur des questions médicales controversées. Il semble de premier abord impossible de savoir, dans l'hypothèse où le CDP du NIH avait un intérêt épistémique, si cela était dû d'une quelconque manière à la participation du public. Je montre au contraire dans la cinquième section de l'article que le CDP du NIH est un cas privilégié pour étudier l'impact d'un forum public sur des questions habituellement propres aux scientifiques et aux chercheurs.

Une attention particulière est portée dans cet article à deux conférences de consensus du CDP administrées par le bureau des applications de la recherche médicale (Office of Medical Applications of Research — OMAR), car elles sont toutes deux utilisées comme exemples par les auteur.e.s précédemment cité.e.s : 1) la conférence sur l'*Helicobacter pylori* dans la maladie ulcéreuse gastroduodénale (NIH, 1994) ; 2) la conférence sur l'élaboration d'un consensus concernant la gestion de l'hépatite C (NIH, 2002). Pour indication, même si le CDP a été supprimé en 2013⁷, des archives demeurent disponibles. Il s'agit surtout d'examiner l'utilité, la légitimité (du point de vue épistémique) et les limites de ces conférences — je laisse de côté les autres considérations développées par Kosolovsky, Van Bouwel et Solomon dans les articles cités ici, telles que, par exemple, celles sur le méta-consensus et celles sur les procédures de délibération (agrégatives ou non). Je défendrai *in fine* une position philosophique intermédiaire sur l'intérêt épistémique des conférences de consensus en question, montrant que les arguments de Solomon, Kosolovsky et Van Bouwel ne sont finalement pas incompatibles, et je montrerai qu'au moins une des deux conférences de consensus analysées a permis, en grande partie grâce à un forum public, d'inclure à la déclaration de consensus des éléments épistémiques et éthiques au préalable passés inaperçus.

2. Distinction entre le consensus académique et le consensus d'interface

L'abondante littérature sur les consensus scientifiques témoigne du fait que l'expression peut faire référence à différents types de consensus scientifiques et de procédures de décision, ainsi que revêtir différentes significations. Communément, le « consensus » peut faire référence de manière générale, d'une part, au résultat visé lors d'une prise de décision collective, à un moment donné et sur un sujet donné, et d'autre part au résultat obtenu ; il peut aussi renvoyer au procédé grâce auquel ce résultat est obtenu, c'est-à-dire faire référence à un système de prise de décision collective (différent, par exemple, d'une décision imposée) ; il peut encore refléter une opinion majoritaire. Selon les auteurs, le résultat peut être obtenu (ou non) par un vote, requiert (ou

⁷ Le programme prend donc fin après la parution des textes de 2007 et 2011 de Solomon, ainsi qu'après la publication de 2012 de Kosolovsky, mais avant l'article de 2014 de Kosolovsky et Van Bouwel.

non) l'unanimité, ou encore, exige que des points de vue minoritaires soient entendus. Aucune définition précise du « consensus » en général, au-delà du sens commun, ne sera proposée dans cet article⁸ ; néanmoins, la distinction entre au moins deux types de consensus liés au domaine scientifique est exposée dans cette section. Au sein du domaine scientifique, c'est la communauté scientifique qui prend part aux décisions et en particulier les scientifiques et chercheurs spécialistes du sujet. Un consensus scientifique est parfois nécessaire pour apporter un jugement consensuel sur une question urgente (ce vaccin est-il efficace ?), ou encore sur un sujet controversé (le vaccin contre la rougeole, la rubéole et les oreillons joue-t-il un rôle déclencheur dans l'apparition des troubles du spectre autistique⁹ ?). Certaines questions relevant du domaine scientifique concernent néanmoins la société et incitent à prendre en compte l'avis du public. L'un des arguments principaux permettant de défendre la participation du public peut alors être résumé de la façon suivante : puisque les citoyens sont concernés, il peut être approprié de les concerter. Il s'agit donc surtout d'un argument éthique et démocratique. Cependant, il n'est pas impossible qu'une participation du public puisse également avoir un intérêt épistémique. Pour le savoir, l'une des solutions envisageables pourrait être de comparer, sur une *même* question, le résultat obtenu au sein d'un consensus établi par des scientifiques spécialistes du sujet au résultat consensuel obtenu lorsque d'autres acteurs ont été conviés à la discussion ou à la délibération¹⁰. Pour cela, dans un premier temps, une distinction peut être utile : la distinction entre ce qui est appelé le « consensus académique » et ce qui est appelé le « consensus d'interface ». Cette distinction est au préalable introduite par Kosolovsky (2012), puis à nouveau utilisée par Kosolovsky et Van Bouwel (2014), et peut également, dans une moindre mesure, être observée dans l'article « Proposal for an Institution for Scientific Judgment » d'Arthur Kantrowitz (1967), bien que l'auteur n'utilise pas directement les intitulés « consensus académique / consensus d'interface ». Cette distinction qui sera utile pour la suite de l'article est présentée ici, mais il s'agit également d'apporter quelques commentaires, permettant notamment d'envisager les impacts de cette distinction, d'une part sur la démarcation entre les discours d'experts et de non-experts, et d'autre part sur l'autonomie des chercheurs.

⁸ À titre indicatif, ma propre conception du consensus rejoint la notion d'acceptation conjointe de Margaret Gilbert dans « Modeling Collective Belief » (1987), selon laquelle « [u]n groupe accepte conjointement p si et seulement si les membres individuels ont ouvertement accepté de laisser p être la position du groupe (Gilbert, 1987, p. 194). [...] [O]n ne parle pas de consensus sur une certaine proposition p , mais de consensus pour faire de p la position du groupe. Sur p lui-même, il pourrait y avoir un désaccord considérable » (cité par Kosolovsky et Van Bouwel, 2014, p. 84 — la présente découpe du texte, de même que les traductions tout au long de l'article, sont les miennes). Kosolovsky et Van Bouwel citent également les travaux de John Beatty et Alfred Moore (Beatty, 2006 ; Beatty et Moore, 2010), qui interviennent plus tard dans cet article, et qui constituent des sources précieuses pour une introduction aux consensus et aux consensus scientifiques.

⁹ Il s'agit, pour rappel, d'une ancienne controverse : les scientifiques s'accordent, au moyen notamment de méta-analyses, sur l'absence de preuve d'un quelconque rôle de la vaccination dans l'apparition des TSA ; voir par exemple à ce sujet les travaux de L.E. Taylor, A.L. Swerdfeger et G.D. Eslick (2014).

¹⁰ Un lecteur avisé ajouterait, en outre, qu'il faudrait que les objectifs visés soient les mêmes pour les deux groupes, et qu'il soit possible de déterminer si les différences qui subsistent dans les résultats obtenus par consensus au sein de deux groupes proviennent directement (ou non) du public, entendu ici comme le grand public ou les citoyens.

Le premier type de consensus mis en avant par Kosolovsky et Van Bouwel est ce qu'ils appellent un « consensus académique, technique » (Kosolovsky et Van Bouwel, 2014, p. 72). Il s'agit d'un « consensus établi par des scientifiques ou des experts d'un certain domaine » (Kosolovsky et Van Bouwel, 2014, p. 72). Le second type de consensus est un consensus établi aux frontières de la science, à l'interface entre science et société. C'est pourquoi les auteurs l'appellent « consensus d'interface ». Ce type de consensus inclut « généralement un plus large éventail d'acteurs en dehors des scientifiques (c'est-à-dire des non-initiés, des experts interactionnels, des représentants du gouvernement, etc.) » (Kosolovsky et Van Bouwel, 2014, p. 72). Ces deux types de consensus sont tous deux liés au domaine scientifique, mais ils impliquent des objectifs différents : au sein du premier, les scientifiques visent un résultat interne à la communauté scientifique et à sa pratique, alors qu'au sein du second, les autres sphères de la société sont intégrées à la discussion et aux délibérations. Ces deux types de consensus impliquent également différents types d'acteurs. Au sein du consensus académique, seuls les scientifiques — qui sont, en outre, spécialistes du sujet abordé — interviennent. D'autre part, selon Kosolovsky et Van Bouwel, les acteurs qui interviennent dans la prise de décision sont sur un même pied d'égalité. En effet, « dans le monde académique, chaque scientifique/universitaire est considéré comme un pair (égal) et chacun fait autorité dans son domaine » (Kosolovsky et Van Bouwel, 2014, p. 73). Puisque Kosolovsky et Van Bouwel incluent aux consensus d'interface les « non-initiés », ce type consensus peut donner une place, en plus des scientifiques et experts, à des scientifiques ou chercheurs non-spécialistes du sujet abordé, des groupes d'intérêts, des citoyens ou leurs représentants, des patients, etc. Avant d'aller plus loin, il est déjà possible de faire plusieurs remarques¹¹ sur cette distinction.

Premièrement, la distinction témoigne du fait que l'on assiste à notre époque à une redéfinition de la démarcation traditionnelle entre science et public du point de vue de l'expertise, ou plus précisément à la redéfinition des limites entre scientifiques et non-scientifiques, experts et non-experts. En effet, au sein du consensus académique tel que défini par Kosolovsky et Van Bouwel, ce n'est pas exactement l'avis du scientifique qui compte, mais celui du scientifique spécialiste du sujet concerné : « ces personnes sont généralement considérées comme à la pointe de la recherche et sont censées être parmi les premières à remarquer les changements survenus dans leur domaine d'expertise » (Kosolovsky et Van Bouwel, 2014, p. 73). Dans le second groupe, au sein des consensus d'interface, il n'y a pas seulement le public ; il peut y avoir des scientifiques spécialistes du sujet abordé et des non-spécialistes, ainsi que des experts non scientifiques tels que les patients experts. Ceci a un impact sur la façon dont est interprétée l'expertise. L'une des premières façons de concevoir celle-ci, lorsque les scientifiques et le public sont distingués, était de situer l'expertise chez les premiers. La distinction de Kosolovsky et Van Bouwel montre que la notion d'expertise devient aujourd'hui plus complexe : il existe une séparation entre les experts spécialistes, certifiés ou non, et les non-spécialistes, certifiés ou non¹². En

¹¹ Les remarques qui suivent ne sont pas abordées par Kosolovsky et Van Bouwel (tout du moins dans les articles de 2012 et 2014).

¹² Voir pour approfondir ce sujet : « The Third Wave of Science Studies: Studies of Expertise and Experience » (Collins et Evans, 2006).

effet, les scientifiques (certifiés) non-spécialistes n'interviennent pas dans le consensus académique. Ils interviennent en revanche dans le consensus d'interface aux côtés d'experts non certifiés. La place des experts non certifiés demeure quant à elle ambiguë : d'un côté, ils peuvent être spécialistes du sujet discuté, mais d'un autre côté ils restent exclus du consensus académique. Si cette place peut être justifiée, l'apparition, en outre, de nouveaux diplômes tels que ceux obtenus au sein de l'Université des patients de la Sorbonne Université (cf. Tourette-Turgis, 2013) crée ce qui pourrait être appelé des « patients experts certifiés » pouvant à nouveau bousculer cette répartition. En effet, ces patients experts *certifiés*, d'autant plus lorsqu'ils sont formés sur des questions techniques (médico-scientifiques), pourraient susciter de nouveaux ajustements concernant la distinction entre consensus académique et consensus d'interface. En plus d'apporter des savoirs expérimentiels sur la maladie, la vie avec la maladie, les traitements et les symptômes, ces types d'acteurs ayant fondé une expertise particulière pourraient devenir de plus en plus légitimes, que ce soit pour recueillir des informations ou pour former les médecins à la prise en charge ou au sujet de maladies rares pour lesquelles les connaissances sont limitées.

Deuxièmement, la distinction entre consensus académique et consensus d'interface peut également avoir un impact sur l'autonomie du chercheur, son interprétation et ses modalités. La réticence vis-à-vis d'une participation du public sur des questions scientifiques peut s'expliquer par le fait qu'une telle participation peut nuire à l'autonomie de la recherche. Pour certains, et à juste titre, la communauté scientifique s'autorégule et devrait par conséquent être la seule à décider des priorités de la recherche. Suivant la distinction de Kosolovsky et Van Bouwel, les chercheurs conservent leur pleine autonomie au sein du consensus académique : ils sont seuls à décider. Au sein cependant des consensus à l'interface entre science et société, on peut supposer que leur autonomie peut être menacée. Cela dépend du rôle accordé aux non-scientifiques et aux non-spécialistes au sein du consensus d'interface. Si le consensus d'interface a un rôle consultatif, les chercheurs peuvent, ou non, prendre en compte le résultat de ce consensus ; du point de vue de la recherche, ils restent maîtres de la décision. Si à l'inverse les autres acteurs du consensus d'interface possèdent un rôle participatif et pas seulement consultatif, l'autonomie de la recherche peut être bousculée.

Ma troisième remarque porte sur les raisons pour lesquelles il faut adopter cette distinction. Kosolovsky et Van Bouwel proposent plusieurs façons d'entrevoir son importance. L'une d'elles est que ces deux types de consensus correspondent à différents moments de la prise de décision : pour qu'il existe un consensus d'interface, il faut qu'il y ait au préalable un consensus académique. D'un point de vue analytique, les auteurs remarquent que ces deux moments méritent alors d'être distingués. Mais ce qui semble également important pour Kosolovsky et Van Bouwel, ce sont les relations qui existent entre les acteurs de ces différents groupes. Comme dit précédemment, au sein du consensus académique, selon eux, les acteurs sont sur un pied d'égalité. Kosolovsky et Van Bouwel insistent sur le fait qu'à l'inverse, au sein du consensus d'interface, le fait qu'il y ait à la fois des scientifiques et des chercheurs et des non-initiés implique une relation « fondée sur l'autorité, la confiance, le respect mutuel, où les acteurs ne sont plus regardés comme étant sur un même pied (épistémique) d'égalité » (Kosolovsky et Van Bouwel, 2014, p. 73). Cet extrait peut,

selon moi, être confus : s'il est vrai que les acteurs au sein du consensus d'interface ne sont pas à égalité du point de vue épistémique — car les savoirs et degrés de savoirs sont hétérogènes —, ce n'est pas tellement parce que la relation est fondée sur l'autorité, la confiance et le respect mutuel. Laisser penser que l'autorité, la confiance et le respect mutuel sont des facteurs d'influence présents uniquement au sein du consensus d'interface, qui mélange plusieurs types d'acteurs, et qu'ils ne seraient pas présents au sein des consensus académiques est illusoire. De nombreux auteurs ont déjà remarqué que les valeurs non épistémiques étaient présentes en science, et que les relations entre scientifiques, même entre spécialistes, étaient aussi chargées de valeurs telles que l'autorité, la confiance, etc. L'ouvrage d'Heather Douglas, *Science, Policy, and the Value-Free Ideal* (2009), par exemple, a montré que les décisions scientifiques ne sont pas uniquement le résultat de valeurs épistémiques, mais plutôt le résultat de valeurs éthiques, sociales, et cognitives. De son côté, Philip Kitcher montre dans *The Advancement of Science* (1993) que des forces sociales telles que l'autorité, la confiance, la réputation et les intérêts personnels sont aussi présents dans les groupes de scientifiques.

Dans cet article, les commentaires de Kosolovsky et Van Bouwel sur l'autorité, la confiance et le respect mutuel seront par conséquent laissés de côté, car ils ne semblent pas constituer pas une raison valable pour distinguer le consensus académique du consensus d'interface. Les raisons pour lesquelles nous retenons qu'il faut, dans la suite de l'article, distinguer le consensus académique du consensus d'interface sont surtout le fait que chaque type de consensus s'établit entre des groupes d'acteurs différents et que les objectifs visés sont également différents.

3. Le consensus d'interface peut-il apporter des éléments épistémiques passés inaperçus au moment du consensus académique ? Le point de vue optimiste

Pour illustrer cette distinction, Kosolovsky et Van Bouwel prennent comme exemple les conférences de consensus du NIH. Alors que pour Solomon, ces conférences de consensus ne peuvent pas apporter d'élément épistémique passé inaperçu au moment du consensus académique, Kosolovsky et Van Bouwel défendent l'idée inverse en prenant comme exemples au moins deux conférences de consensus du NIH également analysées par Solomon. Il s'agit dans cette partie, dans un premier temps, de revenir sur le statut de ce type de conférence : le résultat visé au sein de ces conférences de consensus correspond-il à un consensus d'interface ? Dans un deuxième temps, j'analyse le point de vue de Kosolovsky et Van Bouwel — que je qualifierais de « point de vue optimiste » par opposition au point de vue Solomon, qualifié par ces auteurs de « pessimiste » — pour déterminer si ces conférences de consensus apportent ou non des éléments épistémiques non repérés lors du consensus académique.

Parmi les objectifs visés par ces conférences, cités par le NIH, on trouve entre autres l'évaluation dans un forum public des technologies biomédicales, la publication d'une déclaration de consensus destinée au public et aux praticiens, la diffusion des connaissances scientifiques et le dénouement de questions médicales controversées¹³.

¹³ Pour revenir sur les objectifs du CDP, voir par exemple Wortman, Vinokur et Sechrest (1988).

Solomon, Kosolosky et Van Bouwel insistent sur ce dernier point : l'importance d'aboutir à des consensus rationnels sur des sujets de santé controversés. En ce sens, Kosolosky et Van Bouwel défendent l'idée que les conférences de consensus du NIH peuvent contribuer à la fois au consensus académique et au consensus d'interface : « au premier en établissant un consensus au sein de la communauté scientifique et au second en fournissant et en transmettant ce consensus scientifique établi à une communauté plus large (NIH, site web) » (Kosolosky et Van Bouwel, 2014, p. 73–74). J'aimerais lever un doute ici. Suivant la présentation faite par Kosolosky et Van Bouwel, les conditions de possibilité d'un consensus d'interface semblent être les suivantes : i) le consensus a lieu à l'interface entre science et société ; ii) plusieurs types d'acteurs interviennent dans la prise de décision ; iii) pour qu'un consensus d'interface ait lieu, il faut qu'un consensus académique existe au préalable. De ce point de vue, les conférences de consensus du NIH considérées dans cet article ne seront pas simplement envisagées comme « pouvant participer au consensus d'interface » ; elles sont typiquement des lieux où il est possible d'obtenir un consensus d'interface. Les conférences de consensus ont un statut ambigu : elles réunissaient au départ des acteurs du monde académique, mais sont progressivement devenues un moyen d'inclure les citoyens aux délibérations, sur des questions d'ordre national ou international (pouvant être liées au domaine sanitaire, au domaine environnemental, à un sujet controversé, ou encore à l'utilisation de nouvelles technologies).

Voici une courte présentation de ce que sont en particulier les conférences de consensus du NIH :

La conception de base d'une conférence de consensus du NIH est la suivante. OMAR, un institut ou un centre du NIH, une autre agence gouvernementale de santé, le congrès, ou le public suggèrent des sujets. OMAR décide des sujets à poursuivre. Un comité d'organisation (composé d'employés fédéraux de l'OMAR, de l'institut ou du centre du NIH concerné, d'experts extérieurs, de représentants de patients et du président du comité présélectionné) se réunissent pour choisir les questions qui encadreront la conférence. Les membres du panel pour la conférence de consensus sont choisis parmi les cliniciens, les chercheurs, les méthodologistes et le grand public [...]. Les conférences de consensus ont toujours été ouvertes au public et sont maintenant diffusées par webcam. Les membres du panel et l'auditoire écoutent les présentations académiques de 20-30 experts sur les questions débattues, et ont l'opportunité de poser des questions (Solomon, 2007, p. 167–168).

Plusieurs types d'acteurs participent donc bien à la discussion et aux prises de décision dans les conférences de consensus du NIH. S'ajoute à cela, comme le remarquent Kosolosky et Van Bouwel, que sont distinguées les conférences de développement de consensus (Consensus Development Conference) — qui sont concernées ici — et les conférences ayant pour vocation de faire état de la recherche (State-of-the-Science Conferences). Les premières ont lieu lorsqu'un consensus scientifique est déjà établi, puisqu'elles sont « organisées lorsqu'il existe un ensemble solide de preuves de haute qualité, telles que des essais randomisés et des études

d'observation bien conçues » (Kosolosky et Van Bouwel, 2014, p. 73), alors que les secondes¹⁴ « sont généralement utilisées dans les cas où la base de preuves est plus faible » (Kosolosky et Van Bouwel, 2014, p. 73). Les trois conditions pour obtenir un consensus d'interface peuvent donc être satisfaites.

Quoiqu'il en soit, la question qui intéresse Kosolosky, Van Bouwel et Solomon porte surtout sur la question de l'apport épistémique de ce type de conférences. Le point de vue de Solomon peut être qualifié de « pessimiste », bien que, comme je le montrerai plus tard, il soit souvent proche de la réalité. Dans le chapitre intitulé « The Social Epistemology of NIH Consensus Conférences » (2007), puis dans « Group Judgment and the Medical Consensus Conference » (Solomon, 2011), et plus tard dans son livre *Making Medical Knowledge* (Solomon, 2015), l'auteure montre que les conférences de développement de consensus du NIH ne conduisent pas à un consensus rationnel sur des sujets de santé controversés, qu'un consensus académique existe généralement à l'avance, et que la conférence de consensus ne fait que répéter ce qui a déjà été admis au sein du monde académique. Ces conférences permettent selon Solomon de disséminer les connaissances scientifiques auprès du public et de justifier la pratique clinique, mais elles ne permettent pas d'améliorer les résultats scientifiques. Kosolosky et Van Bouwel résument le point de vue de Solomon de cette façon : ces conférences correspondent à « un rituel, une performance épistémique chorégraphique, créant des connaissances faisant autorité » (Kosolosky et Van Bouwel, 2014, p. 74). Autrement dit, ces moments n'apportent pas de nouvelles connaissances et deviennent alors une perte de temps (et d'argent), uniquement présents pour créer un semblant d'autorité utile aux instances décisionnelles et politiques.

Le second point de vue, celui de Kosolosky et Van Bouwel, est plus optimiste. Ils affirment que ce type de conférence « met en évidence des intérêts épistémiques qui n'ont pas été abordés au moment académique » (2014, p. 78). Cela pourrait donc signifier à l'inverse que prendre en considération le point de vue de différents types d'acteurs, dont le public et les non-spécialistes, n'est pas inutile du point de vue épistémique puisqu'ils permettent d'améliorer la recherche en apportant des éléments négligés lors du moment académique. Pour affirmer cela, Kosolosky et Van Bouwel font appel à deux conférences également examinées par Solomon : la conférence sur l'*Helicobacter pylori* dans la maladie ulcéreuse gastroduodénale (NIH, 1994) et la conférence sur l'élaboration d'un consensus concernant la gestion de l'hépatite C (NIH, 2002).

La conférence sur l'*Helicobacter pylori* dans la maladie ulcéreuse gastro duodénale a lieu en 1994. Selon Solomon, la conférence de consensus « a eu lieu après d'importants essais cliniques (à la fin des années 1980 et au début des années 1990, parmi lesquels certains étaient parrainés par le NIH) et après que des chercheurs, scientifiques et de nombreux cliniciens de premier plan soient parvenus à un consensus sur

¹⁴ Le fait qu'un groupe s'accorde à dire qu'un ensemble solide de preuves de haute qualité est disponible, et qu'il est suffisant pour affirmer qu'il y a un consensus, n'est pas si facile à établir ; la question réapparaît dans l'article. Des extraits analysés plus loin montrent que Solomon aurait été d'accord avec cette affirmation, du moins au moment où elle a écrit les œuvres citées dans cet article, mais que Kosolosky et Van Bouwel ne le seraient pas, ou pas tout à fait.

l'utilisation d'antibiotiques pour les ulcères gastroduodénaux » (Solomon, 2007, p. 170). Selon l'auteure, la conférence de consensus ayant lieu en 1994 n'apporte rien de nouveau. Considérons — indépendamment des analyses fournies par les auteur.e.s en question — ce qui est dit après la conférence de consensus de 1994. Le rapport du NIH résume les conclusions de la conférence de consensus de la façon suivante :

Parmi leurs résultats, le panel de consensus a conclu que : (1) les patients atteints d'ulcère avec infection à *H. pylori* nécessitent un traitement avec des agents antimicrobiens en plus des médicaments anti-sécrétoires, que ce soit lors de la première présentation de la maladie ou en cas de récurrence ; (2) l'intérêt du traitement des patients atteints de dyspepsie non ulcéreuse infectés par *H. pylori* reste à déterminer ; et (3) la relation intéressante entre l'infection à *H. pylori* et les cancers gastriques nécessite des recherches plus approfondies (NIH, 1994, p. 1).

Si l'utilisation d'antibiotiques est confirmée pour les ulcères gastroduodénaux (cf. 1.), le rapport de la conférence de consensus avertit des dangers d'un usage abusif de ceux-ci. Il est par exemple précisé dans ce rapport que « le recours généralisé des traitements antimicrobiens pour traiter l'infection à *H. pylori* peut aggraver le problème de la résistance aux médicaments [...] [et que] des stratégies alternatives de traitement ou de prévention telles que les vaccins ou l'immunothérapie pourraient mériter une attention à l'avenir » (NIH, 1994, p. 9). De plus, le rapport préconise d'utiliser des traitements antimicrobiens uniquement pour les patients positifs à *H. pylori* ayant un ulcère gastrique ou duodénal (NIH, 1994, p. 15). Cela signifie que les patients positifs à *H. pylori* asymptomatiques (sans ulcère) et ceux ayant une dyspepsie non ulcéreuse ne devraient pas en bénéficier. C'est pourquoi la conclusion du rapport indique que « l'intérêt du traitement des patients atteints de dyspepsie non ulcéreuse infectés par *H. pylori* reste à déterminer » (voir ci-dessus, point 2). Pour les patients positifs à *H. pylori* asymptomatiques, le rapport mentionne que les données ne sont pas suffisantes pour envisager un traitement antimicrobien ; pour les patients positifs à *H. pylori* ayant une dyspepsie non ulcéreuse, aucune donnée convaincante n'existe (NIH, 1994, p. 14). Le rapport fait état d'appels à approfondir les recherches sur plusieurs sujets : 1) pour les causes pouvant conduire à l'ulcération duodénale (NIH, 1994, p. 6) ; 2) sur le choix de l'agent antimicrobien pour pouvoir déterminer la thérapie la plus optimale (NIH, 1994, p. 8) ; 3) sur les populations qui ont été jusqu'alors peu étudiées, telles que les enfants ou les patients ayant eu des complications (NIH, 1994, p. 16) ; 4) sur les patients infectés par *H. pylori* qui seront les plus susceptibles de développer un cancer gastrique (NIH, 1994, p. 17). En examinant le même rapport, Kosolovsky et Van Bouwel insistent sur ces demandes de recherches nouvelles et concluent : « le moment de l'interface n'est pas seulement un moment de diffusion conviviale, mais aussi de critique et de contestabilité du moment académique » (Kosolovsky et Van Bouwel, 2014, p. 81). Laisser la place à des contestations est en effet essentiel en science. Kosolovsky et Van Bouwel font référence dans l'article de 2014 aux considérations développées par John Beatty dans l'article « Masking Disagreement Among

Experts » et à celles élaborées par Beatty avec Alfred Moore dans « Should We Aim for Consensus? » (Beatty, 2006 ; Beatty et Moore, 2010). Selon ces auteurs, ce qui est vraiment important dans le consensus scientifique ce n'est pas tant l'unanimité, même si une convergence des opinions est nécessaire à un certain point pour prendre une décision, mais la possibilité de contestation. Ce que nous voulons vraiment savoir, c'est s'il y a eu des désaccords, et si les opinions minoritaires furent entendues. C'est grâce à cela (en partie) que les décisions scientifiques font autorité. Pour ce qui est de l'exemple de la conférence de consensus de 1994, il n'est pas certain que l'on puisse parler de contestabilité en ce qui concerne les résultats : l'usage des antibiotiques dans le traitement de la maladie ulcéreuse gastroduodénale ne fut pas réfuté en soi. De plus, il est difficile pour ce cas — contrairement à la conférence de 2002 examinée ci-après — de vérifier s'il y a eu de nouveaux résultats au moment du consensus d'interface, étant donné qu'aucun document cité ne permet de prendre connaissance au préalable de ce qui a été notifié au moment du consensus académique. Pour rappel, Solomon renvoie seulement aux multiples études effectuées dans les années 1980 et 1990. Ce qui demeure valide, par contre, c'est que la conférence de consensus a joué un rôle de test¹⁵, de supervision¹⁶, voire de contestabilité, en interrogeant en particulier le caractère satisfaisant des preuves disponibles et en mettant en évidence les données manquantes. L'exemple de cette conférence montre aussi qu'en comparaison avec le moment académique, une attention particulière fut portée au moment du consensus d'interface à la prise en charge des patients et aux différentes recommandations à adopter suivant les types de patients.

La conférence sur l'élaboration d'un consensus concernant la gestion de l'hépatite C a lieu en 2002. Kosolovsky et Van Bouwel (2014, p. 76) nous apprennent que deux études publiées par la Food and Drug Administration (FDA) en 2001 préconisent l'utilisation en monothérapie de la ribavirine (un antiviral) pour traiter les patients atteints d'hépatite chronique C¹⁷. Solomon (2007, p. 170–173) affirme de son côté que la conférence de consensus, du point de vue épistémique, ne fait que répéter ce qui était déjà annoncé par la FDA l'année précédente. Les nouveautés portent plutôt, selon elle, sur la prévention — au moyen de recommandations pratiques — et sur les sujets à étudier dans le futur — les directions pour la recherche. Pour Kosolovsky et Van Bouwel, à l'inverse, la conférence de consensus de 2002 apporte quelque chose de plus puisque le rapport montre que le traitement thérapeutique le plus efficace pour traiter l'hépatite C chronique n'est pas la monothérapie avec la ribavirine, mais la combinaison de ribavirine et d'interféron pégylé (Kosolovsky et Van Bouwel, 2014, p. 78–79 ; NIH, 2002, p. 17). Cette information semble donner en partie raison à Kosolovsky et Van Bouwel : le CDP offre dans le cas de 2002 des

¹⁵ Les résultats résistent-t-il à l'importation d'un autre système de valeurs, d'un autre pays, d'une autre communauté, d'un autre groupe de scientifiques ?

¹⁶ Inclure un groupe d'acteurs de différents types (dont le public) ne fait pas qu'apporter des valeurs absentes au moment du consensus académique ; le groupe peut aussi servir de corps extérieur en incluant les citoyens, et est alors libre des valeurs propres aux scientifiques (spécialistes du sujet), telles que les intérêts administratifs, économiques ou méthodologiques. Les membres de la conférence de consensus en question n'hésitent pas à demander en effet plus d'études, multipliant les méthodes et les investissements de temps et d'argent.

¹⁷ Voir à ce sujet « Treatment for Hepatitis C » (Schwetz, 2001).

informations différentes de celles présentées dans les recommandations de la FDA. Néanmoins, voici comment il est possible de résumer les recommandations de la FDA parues l'année précédente :

La FDA a délivré deux approbations concernant l'utilisation de gélules de Rebetol (ribavirine), d'injection d'Intron A (interféron alfa-2b) et d'injection de PEG-Intron (peginterféron alfa-2b) (toutes Schering Corp, Kenilworth, NJ) pour traiter les patients atteints d'hépatite C. Les gélules de Rebetol et l'Intron A ont été approuvés en tant que traitement d'association unique pour le traitement des patients atteints d'une infection chronique par l'hépatite C qui n'ont pas reçu de traitement par interféron ou qui ont rechuté après un tel traitement. La nouvelle approbation prévoit la distribution de Rebetol sans Intron A, donnant ainsi aux prescripteurs la flexibilité d'adopter des thérapies individualisées à base de ribavirine et d'interféron (Schwetz, 2001).

De ce point de vue, la FDA ne préconise pas exactement l'utilisation en monothérapie de la ribavirine. D'une part, la nouvelle approbation de la FDA donne aux prescripteurs « la flexibilité d'adopter des thérapies individualisées à base de ribavirine et d'interféron » (ci-dessus). D'autre part, l'efficacité thérapeutique de l'interféron pégylé est déjà reconnue en 2001 (peginterféron alfa-2b, ci-dessus). Cependant, il n'est pas mentionné dans l'extrait ci-dessus que le traitement thérapeutique le plus efficace pour traiter l'hépatite C chronique serait la combinaison de ribavirine et d'interféron pégylé. Cela signifie que l'argument de Kosolosky et Van Bouwel, sans être tout à fait exact, demeure en partie valide : il y a bien une nouvelle information¹⁸. Le rapport de la conférence de consensus de 2002 dit : « Dans l'ensemble, l'interféron pégylé plus ribavirine était plus efficace que l'association interféron-ribavirine standard ou l'interféron pégylé seul » (NIH, 2002, p. 17). En résumé, le rapport de la conférence de consensus de 2002 propose de nouvelles recommandations de politique sanitaire et préventive et introduit une nouvelle information¹⁹ consensuelle (par rapport à celles présentées dans les rapports de la FDA en 2001). Il présente également les résultats d'une « étude récente, encore non publiée » (NIH, 2002, p. 18) sur la combinaison interféron pégylé-ribavirine.

Cette partie de l'article montre que les conférences de consensus du CDP n'ont pas seulement comme objectif de disséminer l'information scientifique auprès du public ou de rendre (éventuellement) la pratique scientifique plus éthique et démocratique. Le premier cas — la conférence sur *Helicobacter pylori* dans la maladie ulcéreuse gastroduodénale de 1994 — a permis d'insister sur les conséquences néfastes d'une

¹⁸ Cela s'explique, comme Kosolosky et Van Bouwel le remarquent, par des questions différentes (bien que proches) posées dans le rapport de la FDA et celui du NIH ; alors que ce qui est en jeu dans le rapport de la FDA est l'efficacité thérapeutique d'un traitement pour l'hépatite C, le rapport du NIH répond à la question : « Quel est le traitement le plus efficace pour l'hépatite C ? » (Kosolosky et Van Bouwel, 2014, p. 79).

¹⁹ Il ne s'agit pas réellement d'une « nouvelle information » tout court, car l'information apparaît déjà dans une étude scientifique au préalable, mais ici d'une « nouvelle information consensuelle », au sens où l'étude est mise en perspective avec d'autres études pour proposer au public et à la communauté scientifique une conclusion.

surmédication, de porter une attention particulière aux recommandations à adopter en fonction des types de patients et surtout d'interroger le caractère satisfaisant²⁰ des preuves disponibles, en mettant de plus en évidence les données manquantes. Le second cas — la conférence sur l'élaboration d'un consensus concernant la gestion de l'hépatite C de 2002 — a permis de contester et/ou d'améliorer les consensus scientifiques préétablis et, là encore, d'identifier les prochaines études à mener, dont les sujets sur lesquels les données sont manquantes ou peu satisfaisantes. Cette partie suggère également que le caractère acceptable d'un résultat obtenu par consensus n'est pas seulement fondé sur la pertinence du résultat, mais aussi sur la procédure par laquelle il fut obtenu ; la confiance qui est accordée à une décision obtenue par consensus dépend en partie des désaccords et des débats qui précèdent la décision finale (Beatty, 2006 ; Beatty et Moore, 2010). La comparaison effectuée dans cette partie semble donc donner raison à Kosolosky et Van Bouwel concernant l'importance de la contestabilité et le fait que les conférences de consensus du NIH (celles-ci au moins) permettaient de discuter notamment le caractère satisfaisant des preuves disponibles. Pour le cas de 2002, ils ont également raison lorsqu'ils affirment, en opposition à Solomon, que le rapport du NIH ne fait pas *que* répéter ce qui était déjà annoncé par la FDA l'année précédant la conférence. Cependant, l'examen de ces deux conférences ne signifie pas que ces conclusions soient généralisables à toute conférence de consensus et ne permettent pas de savoir si les conférences de consensus en question ont permis d'apporter des éléments épistémiques passés inaperçus au moment du consensus académique : pour les deux conférences, le fait qu'il y ait eu au préalable un consensus académique n'est pas clair. Cela ne signifie pas non plus que le point de vue pessimiste de Solomon devrait être abandonné. La partie suivante montre que l'analyse de Solomon sur les conférences de consensus du NIH met en évidence, entre autres, des biais potentiels, la possible manipulation du public et des procédures de délibération inadaptées.

4. Le consensus d'interface peut-il apporter des éléments épistémiques passés inaperçus au moment du consensus académique ? Le point de vue pessimiste

La partie précédente semblait donner raison à Kosolosky et Van Bouwel sur la capacité du Programme de développement de consensus du NIH à remplir l'objectif de servir à la fois le public et le monde académique ; il ne servait pas seulement à disséminer les informations scientifiques auprès du public. Cette nouvelle partie i) revient sur les arguments permettant à Kosolosky et Van Bouwel de critiquer l'analyse de Solomon, et montre qu'ils ne sont pas entièrement valables ; ii) expose d'autres arguments présents chez Solomon (2007, 2011 et 2015) laissant penser qu'il faut plutôt adopter un point de vue pessimiste vis-à-vis des conférences de consensus organisées par le NIH et de leur intérêt épistémique. La partie ultérieure propose une position intermédiaire par la mise en commun des arguments demeurant valides chez ces auteur.e.s.

Selon Kosolosky et Van Bouwel, « Solomon serait d'accord avec le fait qu'une conférence de consensus peut contribuer au consensus d'interface, elle ne serait pas d'accord sur le fait que les conférences de consensus contribuent au consensus

²⁰ Kosolosky et Van Bouwel fournissent une analyse bien plus étayée sur ce sujet ; voir le concept d'« epistemic adequacy » (Kosolosky et Bouwel, 2014, p. 77–81).

académique, car ce dernier consensus existait déjà avant que la conférence commence » (2014, p. 74). En effet, Solomon affirme : « Les conférences de consensus semblent manquer la fenêtre d'opportunité épistémique prévue : elles ont généralement lieu après que les experts aient atteint un consensus » (Solomon, 2007, p. 170). Pour aller à l'encontre de ce qu'on pourrait nommer « l'argument de la fenêtre épistémique », Kosolovsky et Van Bouwel concentrent l'argumentation sur deux aspects : 1) la typicalité présumée des exemples choisis par Solomon ; 2) l'adoption présumée d'un consensus académique au préalable.

Concernant le premier aspect, Kosolovsky et Van Bouwel reprochent à Solomon de ne pas préciser les raisons pour lesquelles les conférences de consensus de 1994 et de 2002 constitueraient des exemples typiques de conférences de consensus du NIH, alors que ces exemples jouent un rôle central dans son argumentation. Selon eux, « La question reste donc de savoir s'ils sont vraiment aussi typiques qu'elle le suppose et sur quelles bases elle (peut) faire cette hypothèse » (Kosolovsky, 2012 ; Kosolovsky et Van Bouwel, 2014, p. 75). Dans son article de 2007, Solomon décrit le contexte dans lequel sont apparues ces conférences, la méthode adoptée et les critères devant être satisfaits pour qu'une conférence du consensus du NIH ait lieu. Il est possible de défendre l'idée que la typicalité des exemples qu'elle choisit ne porte pas tant sur le contenu de ces conférences mais plutôt sur la procédure qui fut mise en place. En ce sens, l'argument de Kosolovsky et Van Bouwel n'a qu'une portée très limitée : les conférences de consensus suivent la même procédure, bien que cette procédure ait pu faire l'objet de quelques modifications au cours du temps. L'Institut national de Médecine (États-Unis) mentionne par exemple :

Selon l'OMAR, quatre critères sont appliqués pour évaluer les sujets potentiels : (1) le problème doit avoir une importance pour la santé publique, (2) il doit y avoir une controverse sur les aspects scientifiques du problème, (3) il doit y avoir des preuves disponibles sur lesquelles baser l'évaluation du problème, et (4) le problème doit pouvoir être clarifié pour des raisons techniques (OMAR, NIH, 1988) (Institute of Medicine, 1990).

De ce point de vue, les conférences de consensus de 2002 et de 1994 sont des exemples typiques car, comme les autres conférences de consensus du NIH, ces conférences ont été planifiées en respectant les critères mentionnés ci-dessus. Les conférences de consensus doivent en particulier avoir lieu lorsqu'il existe « des preuves disponibles sur lesquelles baser l'évaluation du problème » (voir ci-dessus, point 3). S'il n'y a pas au préalable de preuves disponibles ou de preuves suffisantes, le NIH, pour rappel, n'organise pas de conférence de consensus, mais des conférences sur l'état de la recherche (State-of-the-Science Conferences). À chaque fois qu'une conférence de consensus a lieu, des preuves scientifiques sont donc déjà disponibles dans la littérature scientifique.

Cela conduit au deuxième point sur lequel Kosolovsky et Van Bouwel insistent : l'adoption présumée d'un consensus académique au préalable. Tel qu'introduit précédemment, selon Kosolovsky et Van Bouwel, Solomon « n'explique pas comment exactement l'établissement du consensus intra-scientifique et académique s'est fait » (Kosolovsky et Van Bouwel, 2014, p. 75). Suivant ce que défend Solomon, le consensus

scientifique est fondé sur des preuves scientifiques et les résultats de méta-analyses. De ce point de vue, puisque les conférences de consensus du NIH ont toujours lieu lorsque des preuves scientifiques et des méta-analyses sont disponibles, l'auteure défend qu'un consensus entre experts a été atteint. À ce sujet, Solomon cite parallèlement John Ferguson, qui fut pendant longtemps le directeur du programme, et qui écrit en 1993 : « Souvent, les organisateurs d'une conférence de consensus donnée sont conscients du résultat probable et utilisent la conférence comme un mécanisme pour informer la communauté du domaine médical » (Solomon, 2007, p. 170). Kosolosky (2012) remarque à juste titre qu'il s'agit d'un argument d'autorité (sa portée est donc limitée) et qu'un résultat probable ne correspond pas nécessairement au résultat qui sera en effet adopté. La question demeure cependant : les auteurs ne semblent pas adopter la même définition du « consensus académique » et Kosolosky et Van Bouwel s'interrogent sur le point de vue Solomon. La question est de savoir si s'accorder sur le fait qu'il y ait un ensemble de preuves disponibles suffit pour dire, comme le suggère Solomon, qu'il y a un consensus préalablement établi au sein du monde académique (voir Kosolosky, 2012 ; Kosolosky et Van Bouwel, 2014, p. 80). Pour Kosolosky et Van Bouwel, cela ne suffit pas, et Kosolosky insiste sur le fait que le rapport de la FDA ne mentionne pas qu'il y ait au préalable un consensus préétabli au sein du monde académique (2012), alors que Solomon affirme que c'est le cas, pour la conférence de 1994 comme pour celle 2002 ; elle dit plus précisément que ces conférences de consensus « ont généralement lieu après que les experts soient parvenus à un consensus » (Solomon, 2007, p. 170). Ce qui semble donc plus important du point de vue de Solomon est que l'information apparaissait déjà dans la littérature scientifique.

En résumé, selon Kosolosky, puis Kosolosky et Van Bouwel, Solomon défend son point de vue pessimiste sur l'intérêt épistémique du CDP en présentant trois arguments : l'argument de la fenêtre épistémique, l'argument de la « répétition au sein des exemples de 2002 et 1994 des conclusions déjà disponibles », et l'argument d'autorité de Ferguson.

Indépendamment des informations apportées par Kosolosky et Van Bouwel, qui permettent d'apporter des précisions, Solomon propose une multitude d'arguments lui permettant d'exposer les raisons pour lesquelles les conférences de consensus du NIH n'étaient pas constructives ou productrices de connaissances. Je présente ici de façon succincte quelques-unes de ces raisons²¹, en les résumant et sans les développer ; pour faciliter la compréhension, je propose néanmoins des intitulés permettant de percevoir directement le sujet sur lequel porte chacune des réflexions de Solomon.

L'argument de l'objectivité superficielle. Les conférences de consensus du NIH étaient perçues comme étant objectives, car elles étaient conçues de façon à supprimer (qu'elles y réussissent ou non) les biais liés aux pressions politiques et commerciales, ainsi que les biais liés à la recherche antérieure. Cependant, il s'agissait davantage, selon Solomon, de créer une apparence d'objectivité qu'une objectivité réelle (absence de biais et de préjugés). En effet, des évaluations du programme

²¹ Je laisse notamment de côté les arguments de 2015, car l'ouvrage paraît après les publications de Kosolosky et Van Bouwel.

ont montré²² que les conférences de consensus n'étaient pas conçues de façon à éviter les biais liés à la dynamique de groupe, à l'ordre des présentations, à la force rhétorique des locuteurs, à la pression des pairs, à l'évaluation des preuves, et ceux liés de façon générale à la pratique médicale (cf. Solomon, 2007, p. 169–174).

L'argument du temps. Cet argument est double. D'une part, l'organisation de telles conférences demandait environ un an de travail, ce qui signifie qu'entre le moment où le sujet de la conférence de consensus était choisi et le moment où elle avait lieu, un consensus pouvait apparaître au sein de la communauté scientifique et de nouvelles études pouvaient être publiées avec de nouveaux résultats. D'autre part, le temps imparti pour la conférence de consensus — c'est-à-dire deux jours et demi, contrairement au temps imparti dans d'autres pays organisant des conférences de consensus — était trop court (cf. Solomon, 2011, p. 243–246). Cela signifie que les membres n'avaient ni le temps de lire toutes les études, ni le temps de mener une réflexion approfondie ou de collecter des informations supplémentaires.

L'argument des réunions nocturnes. Certaines sessions avaient lieu tard le soir et la rédaction de la déclaration de consensus pouvait prendre la nuit entière, ce qui favorise d'une part l'apparition de biais potentiels, et d'autre part les opinions de ceux qui sont assez robustes pour être encore présents (cf. Solomon, 2011, p. 245–246).

L'argument du manque d'expertise et du manque de public. Bien que les conférences de consensus du NIH accueillent des scientifiques travaillant sur des sujets proches de ceux abordés, elles excluaient ceux qui avaient le plus de connaissances sur ce sujet (ceux ayant, par exemple, publié sur le sujet concerné) dans le but d'éviter des préjugés intellectuels liés à la recherche antérieure (cf. Solomon, 2011, p. 246). D'autre part, si le public était bien invité, d'autres pays invitaient plus de public, dont notamment des journalistes et des représentants politiques, dans le but d'encourager les discussions portant sur l'éthique, l'économie et la politique, que le NIH de son côté tentait d'éviter (cf. Solomon, 2011, p. 246).

L'argument de non-réussite (à produire un consensus inédit). Lorsqu'il n'existait pas au préalable de consensus au sein de la communauté scientifique, les conférences de consensus du NIH n'en produisaient pas ; lorsqu'il n'existait pas au préalable de consensus au sein de la communauté scientifique, le NIH organisait plutôt des conférences sur l'état de la recherche (cf. Solomon, 2007, p. 170).

S'ajoutent à ces arguments deux points, entre autres, sur lesquels Solomon s'interroge : le premier concerne la nécessité d'organiser une conférence de consensus, le second porte sur la pertinence d'un consensus produit par des non-spécialistes.

La nécessité d'organiser une conférence de consensus (au sein du domaine médical). Lorsque dans d'autres domaines — l'auteure cite à titre d'exemples la physique et la géologie — les scientifiques ne sont pas d'accord sur un sujet, ils concentrent leur énergie sur la recherche de preuves supplémentaires, ils n'organisent pas une conférence de consensus (cf. Solomon, 2011, p. 247). Solomon pose la question suivante : « Pour le dire clairement, pourquoi un groupe d'éminents médecins assis dans une salle de conférence d'hôtel serait-il capable de résoudre une controverse entre

²² Solomon cite entre autres un examen interne paru en 1980, un examen de la Rand Corporation publié en 1983, une étude de l'Université du Michigan parue en 1987, une étude de l'Institut de Médecine (États-Unis) réalisée en 1990, et un groupe de travail du NIH réuni en 1999.

chercheurs du domaine médical ? » (Solomon, 2011, p. 247). L'auteure cite le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) à titre d'exception, mais pour ce cas, la création d'une conférence de consensus est appropriée selon l'auteure, car le GIEC ne vise pas à informer la communauté scientifique, mais plutôt les non-scientifiques (public, instances décisionnelles), tout en suggérant des recommandations politiques (et, à nouveau, non scientifiques).

La pertinence d'un consensus produit par des non-spécialistes. S'il y a une controverse scientifique sur un sujet, c'est par les preuves scientifiques qu'elle devrait être résolue : « et si elle ne peut être résolue par le rapport de preuves, elle ne devrait pas être résolue par d'autres moyens, sous peine de rendre le processus non fondé sur des preuves » (Solomon, 2011, p. 249–250) — c'est en tout cas le principe de l'« evidence-based medicine » (EBM). La question « si un panel obtient un rapport de preuves sur le sujet de la conférence de consensus, que leur reste-t-il à faire ? », posée par Solomon (2011, p. 249), témoigne selon moi du fait que pour elle, l'étude formelle des conclusions qui peuvent être tirées de plusieurs études scientifiques *vaut* consensus, ou à tout le moins, que le consensus devrait être fondé sur ce type d'outils et non sur la délibération entre experts et public²³.

Cette partie dans son ensemble témoigne du fait que les conférences de consensus du NIH étaient mal conçues, d'une part à cause de procédures de décision biaisées à de multiples niveaux, et d'autre part parce qu'elles étaient focalisées sur l'aspect épistémique des questions à traiter. La suppression de ce programme au sein du NIH tend à donner raison à l'auteure. De ce point de vue, l'analyse de Solomon semble bien plus *réaliste* que pessimiste, et s'il s'agit bien d'un point de vue pessimiste, il s'apparenterait davantage à un « pessimisme modéré ». Sur la participation de différents types d'acteurs concernant des questions à résoudre se situant à l'interface entre science et société, l'auteure défend à plusieurs moments l'idée qu'une participation peut être appropriée et bénéfique : pour discuter des facteurs contextuels et des conséquences sociales, éthiques, politiques ou financières (telles que le coût d'un traitement dans un pays donné), pour disséminer les informations scientifiques à d'autres publics (le public en général, mais aussi les instances décisionnelles ainsi que les cliniciens) et les rendre plus accessibles, pour rendre les procédures plus transparentes, ou encore, pour prendre en compte les préférences des patients et leurs représentations de la maladie, ainsi que celles des cliniciens (cf. Solomon, 2007 et 2011, p. 246–247). L'auteure examine également les conférences de consensus créées après celles du NIH et dans d'autres pays, notamment en Europe (Royaume-Uni, Danemark, Pays-Bas et Suède). Son analyse tend à conclure que ces autres exemples de conférence de consensus sont plus appropriés parce qu'elles sont conçues de façon à discuter et délibérer sur des questions éthiques et politiques (dont celles liées au domaine médical) et s'écartent à l'inverse d'une poursuite d'objectifs de type épistémique. Il semble que l'un des objectifs de Solomon soit

²³ Résumer l'opinion de Solomon ainsi est vraisemblablement réducteur, au sens où l'auteure est tout à fait au courant que les méta-analyses peuvent elles aussi être sujettes à des biais, et que dans certains cas les scientifiques doivent faire face à des incertitudes. Voir aussi, en particulier sur le premier point, l'ouvrage *Medical nihilism* de Jacob Stegenga (2018). Sur le second, voir la littérature sur le risque inductif (Brown, 2013 ; Douglas, 2000 et 2017 ; Elliott et Richards, 2017 ; Havstad et Brown, 2017 ; John, 2015 ; Stegenga, 2017).

bien plus de déterminer les moments opportuns pour qu'une participation du public puisse être bénéfique, ainsi que les moments où, à l'inverse, elle n'est pas appropriée, que de nier qu'elle puisse avoir un intérêt d'un point de vue général.

5. Position philosophique intermédiaire et perspectives méthodologiques pour évaluer la participation du public

De façon générale, des moments d'interface entre la science et la société, tels qu'au sein de conférences de consensus du NIH qui intégraient différents types acteurs, dont le public, peuvent générer, d'un point de vue hypothétique, des effets pouvant être perçus comme bénéfiques : elles permettent de suggérer de nouvelles pistes de recherche, de supprimer ou contrebalancer des biais propres aux scientifiques, de favoriser l'élaboration de recommandations davantage en adéquation avec la pratique clinique, d'inclure les savoirs expérientiels des patients, de susciter des discussions plus transparentes sur les facteurs contextuels et les conséquences sociales, éthiques, politiques et financières, et d'assurer une meilleure diffusion des résultats de la recherche, lesquels sont rendus, au demeurant, plus accessibles. Cependant, les nombreux biais décrits par Solomon au sein des conférences de consensus du NIH rappellent, malgré le fait qu'elles fussent longtemps réputées comme étant des sources « sûres », que leur format n'était pas approprié pour produire de tels bénéfices. En outre, les retours ici entrepris sur la conférence au sujet de l'*Helicobacter pylori* dans la maladie ulcéreuse gastroduodénale (1994) et la conférence sur l'élaboration d'un consensus concernant la gestion de l'hépatite C (2002) confirment que ces conférences de consensus n'ont pas permis d'apporter des informations totalement inédites du point de vue épistémique, car de telles informations apparaissaient déjà au préalable dans la littérature scientifique (ce qui ne signifie pas que l'inverse ne puisse jamais arriver). Néanmoins, l'argument de Kosolovsky et Van Bouwel consistant à affirmer que les moments d'interface, contrairement aux procédures de vote agrégatives et aux méta-analyses, possèdent l'avantage de laisser place à des contestations et permettent d'interroger le caractère satisfaisant des preuves disponibles, demeure valide. En résumé, à partir des deux études de cas sur lesquelles s'attardent Solomon d'un côté et Kosolovsky et Van Bouwel de l'autre (1994 et 2002), la mise en commun des deux derniers points mentionnés laisse entrevoir une position intermédiaire : les deux conférences de consensus du NIH prises comme objet d'étude montrent que ces conférences de consensus n'ont pas permis d'apporter d'informations scientifiques absentes au préalable de la littérature scientifique, mais que leur format, sans être entièrement satisfaisant, rendait possible la contestation des résultats (ce qui accroît du point de vue procédural la confiance qui est accordée à un résultat obtenu par consensus). Si ces moments de contestation participent d'une certaine manière à l'objectivité scientifique²⁴, ils possèdent en ce sens un intérêt épistémique, et il serait dommage de les supprimer dans des CDP ultérieurs. La position que je défends dans cette section va plus loin : ces conférences de

²⁴ Le point de vue de Kosolovsky et Van Bouwel sur la contestabilité s'accorde avec ceux d'autres auteurs ; voir par exemple à ce sujet les ouvrages *Science as Social Knowledge* de Helen Longino (1990) et ses considérations sur ce qu'elle appelle « transformative criticism », ainsi que *Science, Policy, and the Value-Free Ideal* de Heather Douglas (2009) et ce qu'elle appelle « interactive objectivity ».

consensus n'ont pas permis d'apporter d'informations scientifiques non-présentes au préalable dans la littérature scientifique, mais, celle de 2002 (au moins) a permis d'inclure à la déclaration de consensus finale des éléments épistémiques et éthiques passés inaperçus lors de précédentes décisions consensuelles. La première partie correspond donc à la thèse défendue par Solomon ; la seconde s'inspire de celle défendue par Kosolovsky et Van Bouwel sans être entièrement identique. Pour rappel, Kosolovsky et Van Bouwel défendent la thèse suivante : « le moment d'interface met en évidence des intérêts épistémiques qui n'ont pas été abordés au moment académique » (Kosolovsky et Van Bouwel, 2014, p. 78). Dans la position intermédiaire énoncée ci-dessus, il s'agit plutôt de dire qu'il est possible que certains éléments *épistémiques* et éthiques fussent abordés lors du moment académique, ou encore dans la littérature scientifique, mais qu'ils sont passés *inaperçus* au moment de rédiger la déclaration de consensus. Pour défendre cette thèse et approfondir cet aspect, reprenons le cas de la conférence sur la gestion de l'Hépatite C (2002), et en particulier, ce qu'il s'est passé le troisième jour. Ceci va permettre : i) d'illustrer davantage l'argument de Kosolovsky et Van Bouwel sur la contestabilité²⁵ tout en apportant des indications supplémentaires sur leurs affirmations ; ii) de montrer que le programme de développement de consensus du NIH (CDP) est, parmi les autres programmes qui lui ont succédé, l'un des rares cas qui permette d'évaluer l'impact d'un forum public lors de telles conférences.

Pour rappel, les conférences de consensus du NIH duraient deux jours et demi. Plus précisément, pour cette conférence de consensus de 2002, une revue systématique de la littérature sur le sujet de la conférence est mise à disposition avant que ne débute la conférence, préparée par l'Agence pour la recherche et la qualité des soins de santé (Agency for Healthcare Research and Quality — AHRQ). Durant les deux premiers jours, des scientifiques, experts du sujet, sont invités à présenter les preuves scientifiques disponibles sur différents aspects du thème abordé ; les résumés sont présentés et disponibles dans le programme de l'événement. Ces conférences sont rendues publiques et « le public est encouragé à poser des questions, à faire des commentaires et à fournir des preuves supplémentaires, et le panel est autorisé à interroger les orateurs sur leurs présentations et à demander des éclaircissements si nécessaire » (Wortman, Vinokur et Sechrest, 1988, p. 471). En prenant appui sur ces deux premières journées, les membres du conseil restreint écrivent une première déclaration de consensus (comme précisé précédemment, parfois jusque tard dans la nuit). Le troisième jour, cette déclaration de consensus est entièrement lue et présentée lors d'un forum ouvert où l'auditoire, constitué de tous les membres présents — c'est-à-dire notamment les conférenciers et le public en général —, est invité à proposer des suggestions. La première version est alors, si besoin, modifiée pour tenir compte des suggestions, puis publiée plus tard après révisions. Depuis 1996, une conférence de presse clôt les deux jours et demi des conférences de consensus²⁶. En quoi ces précisions sont-elles importantes à la lumière de la précédente synthèse ? Elles signifient que lorsque Kosolovsky et Van Bouwel soutiennent que les conférences de consensus du NIH

²⁵ Pour rappel, les auteurs se réfèrent également à Beatty (2018) et Moore (Beatty et Moore, 2010).

²⁶ Voir l'article de P.M. Wortman, A. Vinokur et L. Sechrest (1988) pour des informations plus détaillées sur les différentes étapes, y compris ce qui se passe avant et après les conférences.

mettent en évidence des intérêts épistémiques qui n'ont pas été abordés au moment académique, ils font en réalité appel à différents moments sans les distinguer. Concernant en particulier les deux derniers jours, j'appelle ici et dans la suite de la section certains de ces moments T_1 , T_2 et T_3 . T_1 désigne le moment où les membres au préalable désignés (ou membres du conseil restreint) écrivent la première version de la déclaration de consensus ; T_2 correspond au moment où la déclaration de consensus est présentée lors du forum public pour suggestions (soit le troisième jour) ; T_3 est le moment où la première déclaration est éventuellement modifiée pour produire la déclaration finale. En effet, il semble que les auteurs construisent leur argumentation en tenant uniquement compte de la déclaration de consensus finale²⁷, soit de l'aboutissement de T_1 , T_2 et T_3 . Pourtant, ces moments mériteraient d'être distingués, car les acteurs et les objectifs visés ne sont pas les mêmes. Le premier moment exposé ici (T_1), qui n'est pas rendu public, vise à produire un consensus en réunissant, en général, différents types d'acteurs : des chercheurs, des professionnels de santé, parfois des épidémiologistes et des biostatisticiens, et éventuellement des membres provenant d'autres sphères de la société (éthiciens, économistes, avocats, théologiens et usagers)²⁸. Dans le cas de la conférence de développement de consensus de 2002 sur la gestion de l'hépatite C²⁹, il s'agit en particulier de médecins (M.D.) ayant différentes spécialités — maladies digestives, gastroentérologie, oncologie —, ayant par ailleurs différents statuts — doctorants, professeurs, directeurs de service —, et enfin d'un défenseur des droits des patients dans la recherche (Patient Advocates in Research — PAIR). Ceci permet en particulier d'assurer l'inclusivité sociale, c'est-à-dire ici l'inclusion de participants représentant divers disciplines et intérêts³⁰. Autrement dit, à la lumière de la distinction entre consensus académique et consensus d'interface exposée dans la deuxième section de cet article, et de leurs conditions de possibilité énoncées par Kosolovsky et Van Bouwel, il s'agit déjà ici d'un groupe d'interface et non plus d'un consensus académique³¹. Ces mêmes acteurs se réunissent au moment T_3 . Le pouvoir de contestabilité mentionné par les auteurs est donc selon moi particulièrement saillant lors du moment T_2 puisqu'un consensus (d'interface) est déjà établi³² et qu'il est lu lors d'un forum public dont l'objectif est de recueillir d'éventuelles suggestions. À ce moment-là, l'auditoire est généralement composé de 200 à 700 membres tous publics confondus, dont des citoyens non-experts³³. La déclaration finale, dans le cas de la conférence de 2002 sur la gestion de l'hépatite C, mentionne environ 300 membres (NIH, 2002, p. 1).

²⁷ Ils citent le rapport du NIH, soit la déclaration de consensus finale.

²⁸ L'article « Format and Conduct of Consensus Development Conferences: Multination Comparison », utilisé à plusieurs reprises dans cette section, offre une vue d'ensemble des différentes variations présentes dans dix programmes de développement de consensus, dont celle du NIH (McGlynn, Kosecoff et Brook, 1990). Voir en particulier ici le « Tableau 3 » sur la composition des membres du panel.

²⁹ Ces informations sont disponibles dans le document final ou la déclaration de consensus (NIH, 2002).

³⁰ La notion d'inclusivité sociale fait ici référence à ce que décrit Stegenga dans l'article « Three Criteria for Consensus Conferences » (Stegenga, 2016).

³¹ En effet, ce moment vise à produire un consensus (sachant que lorsqu'un consensus scientifique n'est pas au préalable disponible, le NIH organise plutôt des conférences de consensus sur l'état de l'art), et le groupe est composé de différents types d'acteurs, dont un défenseur des droits des patients.

³² Ou, suivant la définition du consensus d'interface adoptée précédemment, il est en train de s'établir.

³³ Voir cette fois-ci le « Tableau 4 » (McGlynn, Kosecoff et Brook, 1990).

La thèse que je défends, étayée ci-après, est alors la suivante : les conférences de consensus du NIH sont un cas privilégié pour étudier l'impact d'un forum public sur des questions usuellement propres aux scientifiques et/ou aux chercheurs (innocuité, efficacité thérapeutique, épidémiologie, recommandations cliniques et orientations de recherche).

Depuis l'an 2000, certains moments des conférences de consensus du NIH sont rendus publics ; les enregistrements vidéo sont accessibles en ligne³⁴. Le public n'a pas accès aux réunions visant à établir la déclaration de consensus puis la déclaration finale (T₁ et T₃), mais il est possible d'observer ce qui se produit lors du forum public, c'est-à-dire au moment T₂. Autrement dit, en rendant les conférences de consensus publiques, le CDP du NIH a rendu possible, d'un point de vue méthodologique, l'évaluation de l'impact d'un forum public sur des questions habituellement réservées aux scientifiques et aux chercheurs. Par exemple, lors du forum public de la conférence de consensus de 2002, plus de 50 interventions³⁵ provenant de l'auditoire ont eu lieu. Celles-ci concernaient, entre autres, la contestation de données présentes dans la première version de la déclaration de consensus (par exemple, des données épidémiologiques au préalable sous-estimées, des données trop vieilles pour être fiables et des données sur les patients issus du milieu carcéral) ; elles comprenaient aussi des questions (sur le dosage approprié pour certains sous-groupes ou sur la pertinence de mentionner la résistance au traitement) ; elles mentionnaient des données manquantes pouvant avoir des conséquences sur la prise en charge (concernant par exemple les guérisons spontanées, ou encore les complications dues aux problèmes de santé mentale) ainsi que des références manquantes (par exemple sur les cas asymptomatiques ou sur le traitement par interféron) ; elles permettaient l'ajout de commentaires (sur le partage des aiguilles, la vaccination, le dépistage, des effets secondaires et des complications, la transplantation), des commentaires sur les choix terminologiques et leurs conséquences (par exemple, « l'âge lui-même n'est pas un risque de progression, c'est l'âge au moment de l'infection », ou sur la distinction entre la consommation de drogue ponctuelle ou quotidienne des patients, ou encore l'impact des termes utilisés, *in fine*, sur les consignes de remboursement), des demandes de correction (sur des études citées, par exemple, alors que d'autres ne l'étaient pas). Cette liste n'a pas prétention à être exhaustive. Bien qu'il soit possible d'identifier l'identité et le statut des membres du conseil restreint, cela est plus difficile pour les personnes qui interviennent dans le « public », souvent mentionnés mais parfois inaudibles. Cependant, il est possible de dire que parmi la cinquantaine d'interventions, certaines sont faites par les mêmes intervenants (certains interviennent plusieurs fois), qu'un nombre conséquent des intervenants proviennent du milieu universitaire (il semble que certains fassent également partie des conférenciers invités à parler les jours précédents), qu'au moins un intervenant vient d'un institut clinique et qu'un autre travaille dans un cabinet privé. Bien qu'une directrice de la communication commence ensuite la

³⁴ Voir : <https://consensus.nih.gov/historical.htm> (NIH, site web).

³⁵ Pour être exact, il y en eut environ 56 ; certaines interventions n'ont pas été comptées (lorsqu'une personne se trompe de question, par exemple, ou alors lorsqu'elles prennent la parole mais la laissent finalement à quelqu'un d'autre sans avoir fait de suggestion particulière).

conférence de presse en disant « qu'un débat public a eu lieu au cours duquel les membres du grand public ont eu l'occasion de poser des questions au panel et de faire des commentaires », les citoyens et usagers semblent grandement absents au moment T₂ : une personne intervient néanmoins en tant que représentante de patients pour suggérer de proscrire la consommation d'alcool pour les patients atteints d'hépatite C tant que les données ne permettent pas de savoir précisément la quantité d'alcool acceptable sans avoir un impact sur la progression de la maladie. Les chercheurs faisant « la queue » pour poser leurs questions ou intervenir, et n'ayant pas tous le temps d'intervenir, le dispositif ne semble pas faciliter la participation de citoyens. Ce bref aperçu montre les limites du dispositif, mais illustre aussi les propos de Kosolovsky et Van Bouwel sur la possibilité de contestation et permet d'apporter une réponse à la question de Solomon précédemment citée : « Si un panel obtient un rapport de preuves sur le sujet de la conférence de consensus, que leur reste-t-il à faire ? ». Sans ce forum public, cette cinquantaine d'interventions, dont certaines furent retenues pour la déclaration de consensus finale, n'aurait pas eu lieu (remise en cause de données quantitatives, validité des études qui sont produites, hiérarchisation des sources scientifiques, demandes de précaution, etc.).

Comme précisé précédemment, même s'il fut abandonné, le Programme de développement de consensus du NIH (CDP) a servi de modèle, national et international, pour l'élaboration de programmes similaires³⁶. Les programmes qui ont succédé comportent différentes variations, pouvant par exemple porter sur l'audience visée, les types de questions abordées lors de la conférence, la définition du « consensus », les acteurs qui sélectionnent le thème à aborder, le temps de préparation avant la conférence de consensus et le groupe qui s'en occupe, la littérature scientifique et les types de données disponibles pour la conférence, le nombre de personnes désignées pour écrire la déclaration de consensus, la participation du public dans la décision finale, la procédure de vote et la mention (ou non) des désaccords et minorités entendues dans la déclaration finale. Une étude comparative menée sur dix programmes de développement de consensus provenant de différents pays — Canada, Danemark, Finlande, Pays-Bas, Norvège, Suède, Suisse, Royaume-Uni et États-Unis — propose une vue d'ensemble de ces différentes variations (McGlynn, Kosecoff et Brook, 1990). Il est possible d'observer au sein de cette étude que seul le CDP du NIH, parmi ces dix programmes, a mis en place ce type de forum public *après* que la (première) déclaration de consensus est établie, en ayant comme objectif de récolter des suggestions d'amélioration pour produire la déclaration finale³⁷. De ce fait, le Programme de développement de consensus du NIH (CDP) possède un intérêt méthodologique particulier dans la perspective d'une évaluation de l'impact de divers types de publics sur des questions se situant à l'interface entre science et société. Le CDP du NIH est, parmi les autres programmes qui lui ont succédé, l'un des

³⁶ Pour les exemples européens, l'ouvrage *Public Participation in Science: The Role of Consensus Conferences in Europe* de Simon Joss et John Durant (Joss et Durant, 1995) peut constituer une bonne introduction.

³⁷ Voir le « Tableau 4 ». Selon cette étude, le programme de développement de consensus développé au Pays-Bas demande également aux personnes présentes (en général, entre 150 et 1000 personnes) si elles sont d'accord avec la déclaration de consensus produite par les membres du conseil. Cependant, il n'est pas notifié que ce moment laisse place aux suggestions d'amélioration ou de correction.

rare cas³⁸ qui permette d'évaluer l'impact d'un forum public lors de telles conférences, et en particulier l'impact d'une participation de divers types de publics, dont les citoyens, sur les recommandations destinées aux praticiens et au grand public. De façon plus générale, le visionnage du troisième jour permet également d'obtenir une vue partielle de la première version de la déclaration de consensus, puisque celle-ci est entièrement lue durant la séance publique ; on apprend également que le moment T_1 s'est terminé vers 3h30 du matin. Autrement dit, l'ensemble des documents disponibles au sujet de ces conférences de consensus du NIH pourrait permettre d'évaluer les types de questions posées par le public lors des deux premiers jours, les types d'arguments, de recommandations et demandes de corrections ayant lieu lors du troisième jour, les types de publics (dont hypothétiquement les citoyens) à l'origine de ces questions et recommandations, et, partiellement, les différences qui subsistent entre la déclaration produite par le groupe d'interface (T_1) et celle produite après le forum public (T_3). Cela permettrait également de mettre (éventuellement) en évidence la présence d'opinions minoritaires sur les questions abordées et de vérifier au cas par cas si la thèse de Solomon — selon laquelle les conférences de consensus du CDP n'ont pas été productrices de connaissances encore non présentes dans la littérature scientifique — demeure valide. En France, les bases méthodologiques préconisées pour l'organisation des conférences de consensus dans le domaine médical (Agence nationale d'accréditation, 1999), ont également un format similaire, mais comme pour les autres programmes internationaux cités ici, le format du troisième jour n'a pas été repris : la séance publique a lieu *avant* la rédaction des recommandations.

Quelques éléments ressortent en définitive de cette analyse. Sur les articles de Kosolosky et Van Bouwel (Kosolosky, 2012 ; Kosolosky et Van Bouwel, 2014), présentés comme étant en opposition avec la position pessimiste de Solomon au sujet des conférences de consensus du NIH, l'analyse a permis de montrer que leurs arguments sur l'importance de la contestabilité n'étaient pas incompatibles avec le constat pessimiste de Solomon en développant la position intermédiaire selon laquelle les conférences de consensus n'ont pas apporté d'informations scientifiques supplémentaires non-présentes au préalable dans la littérature scientifique, mais que leur format permettait de contester et d'interroger le caractère satisfaisant des preuves disponibles — ce qui, vis-à-vis du résultat obtenu par consensus, peut accroître la confiance du public et des instances décisionnelles. Dans la perspective où ces moments de contestation peuvent participer à l'objectivité scientifique, les conférences de consensus avaient, en ce sens, un intérêt épistémique. La synthèse met également en exergue la difficulté à formuler une définition du « consensus académique », sans pour autant la résoudre. En ce qui concerne le dispositif de participation du public analysé ici — le CDP du NIH — la synthèse, en particulier du côté de Solomon, rappelle quelques-uns des nombreux biais qui ont justifié sa suppression. La dernière section met en évidence une autre limite importante : bien que le troisième jour ait pour objectif de récolter les suggestions du public en général, le cas de 2002 montre que les citoyens ne semblent pas, ou peu, intervenir ; les suggestions de corrections et de modifications, par ailleurs importantes,

³⁸ Parmi les dix programmes mentionnés précédemment, il est le seul.

proviennent en grande majorité des universitaires présents dans l'auditoire. Le nombre de suggestions récoltées — dont celles ayant abouti à une modification de la déclaration de consensus initiale — illustre cependant l'importance de ces moments laissant place aux discussions et contestations. Sur l'impact d'un forum public sur les recommandations transmises *in fine* aux professionnels de santé et au public en général, l'article montre que le CDP du NIH est le seul cas parmi ceux cités ici qui permette de l'évaluer — ce qui en fait un cas privilégié pour étudier de manière rétrospective l'impact d'un forum public sur des questions généralement réservées aux scientifiques et aux chercheurs. Pour la conférence de consensus de 2002, cette évaluation, bien que sommaire, montre que le grand public (citoyens, patients, usagers) demeure relativement absent et que l'organisation du forum ne permettait pas réellement sa participation. L'importance d'un tel moment d'interface entre science et société, laissant place aux débats, aux contestations, aux suggestions d'amélioration, et, parallèlement, le caractère très limité de l'intervention des citoyens lors du forum de 2002 — alors que leur intervention est précisément l'un des objectifs visés —, interroge sur l'organisation de tels forums et ouvre peut-être la voie à des dispositifs de participation du public plus appropriés. En effet, bien que le CDP du NIH ait servi de modèle pour établir des programmes similaires dans d'autres pays, dont la France, ces moments d'échange, qui ont pourtant un intérêt épistémique, n'ont pas été repris dans les programmes mentionnés ici. Cela pourrait être davantage pris en compte dans une perspective de création de nouveaux programmes ou d'amélioration des programmes existants, dans l'hypothèse que de tels dispositifs puissent alors réellement inclure les citoyens sur des questions factuelles, et non proposer un simulacre de « forum public », dans lequel la participation du public est visée mais jamais vraiment obtenue, peut-être faute d'être correctement prévue.

Remerciements. J'aimerais remercier Claire Crignon, Renaud Debailly et Antoine Guillain pour l'organisation des journées d'études lors desquelles une autre version de ce travail fut présentée, Aude Bandini pour son soutien à l'organisation de ce numéro spécial, ainsi que les différentes personnes ayant participé à l'évaluation et à la relecture attentives de l'article. Je remercie également les personnes ayant participé au colloque international « democrasci », organisé par Stéphanie Rupy et Baptiste Bedessem, pour leurs commentaires sur la première version de ce travail, et en particulier David Guston pour ses conseils.

Références bibliographiques

- Agence nationale d'accréditation et d'évaluation en santé (ANAES). (1999). *Les conférences de consensus. Bases méthodologiques pour leur réalisation en France*. <https://www.has-sante.fr/upload/docs/application/pdf/guideCC.pdf>
- Beatty, J. (2006). Masking disagreement among experts. *Episteme*, 3(1–2), 52–67. <https://doi.org/10.3366/epi.2006.3.1-2.52>
- Beatty, J. et Moore, A. (2010). Should we aim for consensus? *Episteme*, 7(3), 198–214. <https://doi.org/10.3366/epi.2010.0203>
- Brown, M.J. (2013). Values in science beyond underdetermination and inductive risk. *Philosophy of Science*, 80(5), 829–839. <https://doi.org/10.1086/673720>
- Collins, H.M. et Evans, R. (2006). The third wave of science studies: Studies of expertise and experience. Dans E. Selinger et R.P. Crease (dir.), *The philosophy of expertise* (p. 39–110). Columbia University Press.
- Debray, S. (2021). *Le rôle des discours non-académiques au sein des consensus scientifiques* [communication lors des journées d'études]. Patients-experts, patients-partenaires : quels dispositifs de participation en

- santé, Sorbonne Université, 25 et 26 mars. <https://humanites-biomedicales.sorbonne-universite.fr/evenements-ihb/patients-experts-patients-partenaires-quels-dispositifs-de-participation-en-sante>
- Douglas, H. (2000). Inductive risk and values in science. *Philosophy of science*, 67(4), 559–579. <https://doi.org/10.1086/392855>
- Douglas, H. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press. <https://doi.org/10.2307/j.ctt6wrc78>
- Douglas, H. (2017). Why inductive risk requires values in science. Dans J. Turri, K.C. Elliott et D. Steel (dir.), *Current controversies in values and science* (p. 81–93). <https://doi.org/10.4324/9781315639420-6>
- Elliott, K.C. et Richards, T. (dir.). (2017). *Exploring inductive risk: Case studies of values in science*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190467715.003.0001>
- Gilbert, M. (1987). Modelling collective belief. *Synthese*, 73(1), 185–204. <https://doi.org/10.1007/BF00485446>
- Havstad, J.C. et Brown, M.J. (2017). Inductive risk, deferred decisions, and climate science advising. Dans K.C. Elliott et T. Richards (dir.), *Exploring inductive risk: Case studies of values in science* (p. 101–123). <https://doi.org/10.1093/acprof:oso/9780190467715.003.0006>
- John, S. (2015). Inductive risk and the contexts of communication. *Synthese*, 192(1), 79–96. <https://doi.org/10.1007/s11229-014-0554-7>
- Joss, S. et Durant, J. (1995). *Public participation in science: The role of consensus conferences in Europe*. NMSI Trading Ltd.
- Kantowitz, A. (1967). Proposal for an institution for scientific judgment. *Science*, 156(3776), 763–764. <https://doi.org/10.1126/science.156.3776.763>
- Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford University Press.
- Kosolovsky, L. (2012). The intended window of epistemic opportunity: A comment on Miriam Solomon. Dans B. Van Kerkhove, T. Libert, G. Vanpaemel et P. Marage (dir.), *Logic, philosophy and history of science in Belgium II*. Koninklijke Vlaamse Academie van België.
- Kosolovsky, L. et Van Bouwel, J. (2014). Explicating ways of consensus-making in science and society: Distinguishing the academic, the interface and the meta-consensus. Dans C. Martini et M. Boumans (dir.), *Experts and consensus in social science* (p. 71–92). Springer. https://doi.org/10.1007/978-3-319-08551-7_4
- Longino, H.E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press. <https://doi.org/10.1515/9780691209753>
- McGlynn, E.A., Kosecoff, J. et Brook, R.H. (1990). Format and conduct of consensus development conferences: Multination comparison. *International journal of technology assessment in health care*, 6(3), 450–469. <https://doi.org/10.1017/S0266462300001045>
- NIH. (1994). *Helicobacter pylori* in peptic ulcer disease. NIH Cons Statement. February 7-9, 1994. *NIH Consensus State Sci Statements*, 12(1), 1–22.
- NIH. (2002). Management of Hepatitis C: 2002. Consensus Conference Statement, June 10–12, 2002. *NIH Consensus State Sci Statements*, 19(3), 1–46.
- Schwetz, B.A. (2001). Treatment for Hepatitis C. *Journal of the American Medical Association*, 286(10), 1166–1166. <https://doi.org/10.1001/jama.286.10.1166-jfd10008-2-1>
- Solomon, M. (2007). The social epistemology of NIH consensus conferences. Dans H. Kincaid et J. McKittrick (dir.), *Establishing medical reality* (p. 167–177). Springer. https://doi.org/10.1007/1-4020-5216-2_12
- Solomon, M. (2011). Group judgment and the medical consensus conference. Dans F. Gifford (dir.), *Philosophy of medicine* (p. 239–254). Elsevier. <https://doi.org/10.1016/B978-0-444-51787-6.50009-X>
- Solomon, M. (2015). *Making medical knowledge*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198732617.001.0001>
- Stegenga, J. (2016). Three criteria for consensus conferences. *Foundations of Science*, 21(1), 35–49. <https://doi.org/10.1007/s10699-014-9374-y>
- Stegenga, J. (2017). *Drug regulation and the inductive risk calculus*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190467715.003.0002>
- Stegenga, J. (2018). *Medical nihilism*. Oxford University Press. <https://doi.org/10.1093/oso/9780198747048.003.0011>

- Taylor, L.E., Swerdfeger, A.L. et Elick, G.D. (2014). Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies. *Vaccine*, 32(29), 3623–3629. <https://doi.org/10.1016/j.vaccine.2014.04.085>
- Tourette-Turgis, C. (2013). L'université des patients : une reconnaissance institutionnelle des savoirs des malades. *Le sujet dans la cité*, 2(4), 173–185. <https://doi.org/10.3917/lsdlc.004.0173>
- Wortman, P.M., Vinokur, A. et Sechrest, L. (1988). Do consensus conferences work? A process evaluation of the NIH consensus development program. *Journal of Health Politics, Policy and Law*, 13(3), 469–498. <https://doi.org/10.1215/03616878-13-3-469>

Comment citer cet article ? Debray, S. (2022). Questions médicales controversées, déclarations de consensus et participation du public : le cas des conférences de consensus du National Institute of Health. *Dialogue* 61(1), 55–81. <https://doi.org/10.1017/S0012217322000257>