Une brève histoire du concept de biais en épidémiologie

Nicolas Brault

▶ **To cite this version:**

Nicolas Brault. Une brève histoire du concept de biais en épidémiologie. Epidemiology and Public Health = Revue d'Epidémiologie et de Santé Publique, 2021, 69 (4), pp.215-223. 10.1016/j.respe.2021.04.134 . hal-03325191

HAL Id: hal-03325191
https://hal.science/hal-03325191

Submitted on 2 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S039876202100211X
Manuscript_c1624fb8c54915600529e1fc54636038

Une brève histoire du concept de biais en épidémiologie

A short history of the concept of bias in epidemiology

N. Brault*

InTerACT UP 2018.C102, Institut Polytechnique UniLaSalle, 19 rue Pierre Waguet, BP 30313, F-60026, Beauvais, France

**Auteur correspondant*

Adresse e-mail : nicolas.brault@unilasalle.fr (N. Brault)

© 2021 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

Abstract

Background Bias is a major methodological issue for epidemiology. However, only a few studies have been dedicated to the past and present formulations of the concept of bias. Moreover, the classical definition of bias as systematic deviation from the truth of results or inferences, definition which can be found in dictionaries of epidemiology, does not seem to either match the way epidemiologists use it in practice, or correspond to the different definitions given throughout its history. It is consequently important to elucidate this paradox.

Methods In this historical and conceptual article, we study the different uses of the word "bias" in epidemiological literature, from classic articles in the 1950's about the link between smoking and lung cancer to the most recent epidemiology textbooks, the objective being to analyze the ways in which epidemiologists have defined, applied and modified this concept over time.

Results We show that D.L. Sackett's article on bias in analytic research, published in 1979, put an end, at least temporarily, to reflection in populational epidemiology that started thirty years before. More precisely, we show that Sackett's definition of bias corresponds more to the needs and goals of clinical epidemiology than to those of populational epidemiology. Concomitantly, populational epidemiologists such as K.J. Rothman redefined bias as a threat to the internal validity of a study, and epidemiological study as an "exercise in measurement of an effect rather than as a criterion-guided process for deciding whether an effect is present or not".

Conclusion It is thereby important to draw a distinction between two notions pertaining to bias: an epidemiological concept of bias, viewed as the lack of internal validity of an observational study; and a medical concept of bias, defined as deviation from the truth. The former concerns the design and methodology of epidemiological studies; the latter is more general and impels epidemiologists and physicians to be skeptical, and even critical, towards their own inferences.

Keywords History of medicine; Bias; Randomization; Validity; Research design; Observational study; Evidence-Based Medicine

Résumé

Position du problème La question des biais constitue un enjeu méthodologique central pour l'épidémiologie. Cependant, l'histoire de ce concept n'a fait l'objet que de rares études. De plus, la définition classique du biais comme déviation systématique des

résultats ou des inférences par rapport à la vérité, telle qu'elle peut apparaître dans les dictionnaires d'épidémiologie, ne semble correspondre ni à la manière dont les épidémiologistes l'utilisent en pratique, ni aux différentes définitions qui jalonnent l'histoire de l'épidémiologie. Il s'agit donc d'expliquer ce paradoxe.

Méthodes Cet article historique et épistémologique propose d'analyser, depuis les articles classiques des années 1950 sur le lien entre tabagisme et cancer du poumon jusqu'aux manuels d'épidémiologie les plus récents, les diverses occurrences du mot « biais » ainsi que la manière dont les épidémiologistes utilisent, définissent et rectifient ce concept au fil du temps.

Résultats Nous montrons que l'article de D.L. Sackett de 1979, consacré aux biais dans la recherche étiologique, vient clore, au moins provisoirement, une réflexion amorcée trente années auparavant en épidémiologie populationnelle. Plus précisément, nous montrons que la définition donnée par Sackett du biais se construit moins en relation avec les besoins et les buts de l'épidémiologie populationnelle que de ceux de l'épidémiologie clinique. Au même moment, le concept de biais est redéfini au sein de l'épidémiologie populationnelle, notamment par K.J. Rohtman, comme une menace à la validité interne d'une étude, dans le cadre d'une redéfinition de l'étude épidémiologique comme exercice de mesure.

Conclusion Ainsi, il convient de distinguer entre deux concepts de biais : un concept épidémiologique de biais, conçu comme un défaut de validité interne d'une étude observationnelle ; et un concept médical de biais, conçu comme un écart par rapport à la vérité. Le premier porte sur le plan et la méthodologie des études épidémiologiques ; le second est plus général et invite les épidémiologistes et les médecins à une certaine forme de scepticisme, voire de criticisme, envers leurs propres inférences.

Mots-clés Histoire de la médecine ; Biais ; Randomisation ; Validité ; Plan d'étude ; Etude observationnelle ; Médecine factuelle

1. Introduction

Pour qui s'intéresse à l'histoire du concept de biais au sein de l'épidémiologie, un article apparait comme une référence incontournable et fait date, au sens où il y un avant et un après : il s'agit de l'article de D.L. Sackett [1], intitulé « Bias in Analytic Research », publié en 1979, et cité plus de 2400 fois depuis sa parution. Cet article constitue aujourd'hui un classique. Ainsi, lorsque le *Dictionary of Epidemiology* [2] définit le concept de biais comme une «déviation systématique des résultats ou des inférences par rapport à la vérité, ou les processus conduisant à une telle déviation», il reprend largement la définition donnée par Sackett dans son article : «Tout processus à toute étape de l'inférence qui tend à produire des résultats ou des conclusions qui diffèrent systématiquement de la vérité» [1]. Sackett, d'ailleurs s'inspire lui-même de celle donnée par E.A. Murphy en 1976, qui définit le biais comme un « processus qui à toute étape du raisonnement tend à produire des résultats qui s'écartent systématiquement des vraies valeurs » [3]. Ces références à la vérité peuvent surprendre : les « vraies valeurs » renvoient vraisemblablement à des valeurs mesurées, comme par exemple le risque relatif en épidémiologie. Quant aux déviations des résultats ou des inférences par rapport à la vérité, il faut sans doute comprendre ici qu'un biais conduirait à effectuer une inférence qui mènerait à une conclusion qui serait fausse : par exemple, on pourrait conclure à l'existence d'une association ou d'un lien causal entre un facteur de risque F et une maladie M alors qu'il n'y en pas, ou inversement conclure qu'il n'y a pas d'association ou de causalité entre F et M alors qu'il y en effectivement une. Cet article est aussi connu pour être le premier à proposer un catalogue qui liste 35 types de biais dans les enquêtes étiologiques d'observation : Sackett fait d'ailleurs du «développement continu d'un catalogue annoté de biais» la «priorité de recherche numéro un» pour l'épidémiologie. Il semble d'ailleurs que cette longue liste de biais obère toute possibilité de qualifier une étude épidémiologique de valide : nous verrons que ce n'est pas nécessairement le cas, à condition de passer, comme K. Rothman [4], d'une perspective qualitative à une perspective quantitative de la notion de biais. Enfin, la dernière grande nouveauté introduite par Sackett dans cet article consiste à illustrer «l'amplitude et la direction» des différents biais, ainsi que la «description de mesures préventives appropriées». Cet article de Sackett inaugure donc ce que P. Vinéis [5] a appelé la période des catalogues et des taxinomies de biais, qui précède la période des définitions plus formelles des différents biais, notamment par O. Miettinen ou K.J. Rothman.

Ainsi, à lire l'article de Vinéis consacré à l'histoire de la notion de biais, l'article de Sackett s'inscrirait dans la continuité d'une réflexion sur les biais qui aurait commencé dans les années 1950 avec l'introduction de l'essai clinique randomisé par A. B. Hill [6] [7]. Or cette continuité n'est qu'apparente : en effet, la définition donnée par Sackett constitue une rupture par rapport aux conceptions précédentes, plus précisément une rupture au sein de l'épidémiologie, que Sackett assumera plus tard [8], entre l'épidémiologie « classique » ou populationnelle et l'épidémiologie clinique, qui apparait à la fin des années 1960 [9]. En effet, l'épidémiologie clinique s'intéresse non à des populations, mais bien à des individus : il s'agit pour les épidémiologistes cliniques de fonder leur pratique clinique, donc au chevet du patient, sur les résultats des études considérées comme les plus valides, à savoir les essais cliniques randomisés. Or, en confondant populations et individus, la question de savoir si l'épidémiologie clinique relève encore de l'épidémiologie, par définition orientée vers les populations, mérite d'être posée. La même remarque peut d'ailleurs s'appliquer à l'épidémiologie des facteurs de risque qui est devenue, depuis son apparition dans les années 1950 [10], de plus en plus centrée sur le risque individuel, avec les problèmes épistémologiques mais aussi pratiques, notamment vis-à-vis des politiques de prévention, que cela soulève : B. Rockhill [11] critique ainsi la « privatisation » ou d'une « individualisation du risque » qui peut augmenter les inégalités sociales de santé au lieu de les réduire.

En d'autres termes, notre hypothèse est que la définition donnée par Sackett du concept de biais en 1979, alors même qu'elle est considérée comme la définition canonique du biais, ne correspond pas tout à fait à ce que les épidémiologistes entendaient depuis les années 1950 par le terme de biais. Plus précisément il s'agit de savoir comment cette définition est devenue la définition de référence, alors même qu'elle est pensée moins en relation avec les besoins et les buts de l'épidémiologie populationnelle que de ceux de l'épidémiologie clinique, au sens où ce qui intéresse Sackett n'est pas essentiellement la question de la scientificité de l'épidémiologie, mais plutôt celle de la médecine clinique au chevet du patient. Autrement dit, il s'agit de savoir quelle est la « fonction d'opérateur » du concept de biais en épidémiologie, sa « valeur cognitive », c'est-à-dire la « la possibilité qu'il offre de développement et de progrès » [12] pour l'épidémiologie. Notre thèse est que la principale fonction du concept de biais des années 1950 aux années 1990 a été de critiquer la validité des études épidémiologiques observationnelles (études cas-témoin et études de cohorte), et plus précisément leur méthodologie ou leur design, ce qui a conduit les épidémiologistes à affiner et à faire progresser ces méthodes et design, au point

d'oublier le but premier de l'épidémiologie, à savoir la santé des populations. Le même concept a aussi permis aux épidémiologistes cliniques d'établir une hiérarchie des différents types d'études épidémiologiques, en relation avec des questions thérapeutiques, afin précisément d'éviter les erreurs dans leur pratique clinique quotidienne.

Pour ce faire, nous étudierons tout d'abord le contexte historique de la publication de l'article de Sackett, spécifiquement la conférence des Bermudes consacrée à la controverse sur la méthodologie des études cas-témoin dans laquelle Sackett intervient ; puis nous remonterons le temps pour analyser les « importantes discussions » sur les biais qui eurent lieu au sein de l'épidémiologie, à partir des années 1950, dans le cadre de la controverse sur le tabagisme et le cancer du poumon. Ceci nous permettra de distinguer deux concepts de biais qui sont à l'œuvre actuellement dans les sciences médicales, ou plutôt deux fonctions opératoires de ce concept.

2. La « Conférence de paix des Bermudes » et la première définition du concept de biais

2.1. Le concept de biais au cœur de la controverse

En avril 1978, a lieu aux Bermudes un symposium consacré aux études cas-témoins en épidémiologie, et d'où sera tiré un ouvrage, intitulé « *L'étude cas-témoins : consensus et controverses* » [13]. Cet ouvrage constitue un document historique extrêmement intéressant, bien que peu étudié, pour l'histoire de l'épidémiologie en général, et pour celle du concept de biais en particulier. Le premier point intéressant est l'omniprésence du « biais » dans les discussions : nous dénombrons ainsi 490 occurrences du mot « *bias* » sur les 151 pages du document (en incluant la page de titre et l'index). A titre de comparaison, le syntagme « *case-control* » ne compte que 480 occurrences, alors même qu'il constitue le thème principal de l'ouvrage. Si l'on enlève l'article de Sackett [1], d'où il est extrait, consacré explicitement au concept de biais, il reste tout de même 328 occurrences du biais dans l'ouvrage. Dans l'histoire du concept de biais, cette conférence constitue indubitablement, comme nous allons le voir, un tournant.

Le deuxième point intéressant est que c'est à l'occasion de cette conférence qu'est proposée par M.A. Ibrahim et W.O. Spitzer [14] – et à notre connaissance, pour la première fois dans l'histoire de l'épidémiologie – la classification des biais en trois grandes catégories : le biais de sélection, le biais d'information et le biais de confusion.

Cette catégorisation est d'ailleurs reprise telle quelle par D.G. Kleinbaum, L.L. Kupper, et H. Morgenstern en 1982 [15], qui font explicitement référence à l'ouvrage d'Ibrahim, ainsi que par K.J. Rothman et S. Greenland en 1986 [16]. Si les biais de sélection et d'information sont connus et étudiés depuis au moins les années 1950, le coup de force théorique d'Ibrahim et Spitzer consiste à faire de la confusion une catégorie de biais à part entière, alors même que la question de la confusion était en général traitée à part de celle du biais. Cette précision historique n'ayant jamais été faite à notre connaissance, nous nous permettons donc de le faire.

2.2. Une brève histoire des études cas-témoins

Toujours dans une perspective historique, il est essentiel de retracer brièvement l'histoire des études cas-témoins, qui permet d'expliquer la tenue de cette conférence. Cette histoire ayant déjà été largement faite [17,18], y compris dans le cadre de ce symposium aux Bermudes [19,20], il s'agit de se concentrer sur les éléments de cette histoire pertinents pour notre propos. Tout d'abord, le principe de l'étude cas-témoins est de comparer, de façon rétrospective, « un groupe avec un autre en fonction d'une ou de plusieurs caractéristiques d'intérêt ». Les historiens de l'épidémiologie s'accordent à dire que les premières études cas-témoins, bien qu'elles ne soient pas nommées comme telles, remontent au 19ème siècle, voire avant [17,19], mais que la première étude cas-témoin moderne est l'œuvre de J. Lane-Claypon [18,20] dans une étude sur le cancer du sein [17,19]. Les études cas-témoins qui sont considérées comme fondatrices au plan méthodologique sont néanmoins les quatre études qui paraissent en 1950 et qui ont pour thème la relation entre tabagisme et cancer du poumon, notamment l'étude de R. Doll et A.B. Hill [21]. Celle-ci fait rapidement office de « modèle »[18] des études cas-témoins, par son attention et son soin apportés à la méthodologie. Ces études, par leur objet, traduisent la transition épidémiologique [22] que connaissent les pays occidentaux : il s'agit désormais, pour l'épidémiologie des années 1950, d'étudier les maladies chroniques et leurs facteurs de risque et non plus les maladies infectieuses.

Néanmoins, les résultats de ces études sont rapidement critiqués notamment par R. Fisher [23] et J. Berkson [24,25], au motif qu'elles ne permettraient pas de prouver ce qu'elles prétendent prouver, c'est-à-dire ni une relation de causalité entre le tabagisme et le cancer du poumon, ni d'ailleurs une association statistique entre les deux. Ces critiques, assez féroces (Fisher parle de « propagande terrifiante » [23] contre le tabac), ouvrent une discussion extrêmement intense entre les épidémiologistes, autant

sur le plan du design des études cas-témoins ou de cohorte, que sur les critères de causalité : le concept de biais joue ici un rôle important. Ainsi, en 1959 paraissent deux articles décisifs dans l'histoire de l'épidémiologie : celui qui réunit J. Cornfield, W. Haenszel, E.C. Hammond, A.M. Lilienfeld, M.B. Shimkin et E.L. Wynder [26], et celui de N. Mantel, et W. Haenszel [27]. Dans le premier, les auteurs passent en effet en revue les différentes critiques adressées à la relation causale entre tabagisme et cancer du poumon, notamment celles de Berkson et Fisher, et montrent que tous les résultats, issus d'études soit observationnelles soit expérimentales, convergent vers la même conclusion ; mais aussi qu'ils n'ont trouvé aucun agent ou facteur qui soit « aussi fortement associé au cancer du poumon que le tabagisme » [26]. Mais le point le plus important se situe dans les deux appendices à l'article : en effet, pour la première fois dans l'histoire de l'épidémiologie, les auteurs définissent et formulent mathématiquement la notion de « risque relatif », et utilisent cette formule pour justifier dans le corps de l'article, l'utilisation de mesures relatives de risque plutôt que de mesures absolues, notamment en montrant que la mesure du risque relatif permet aussi de traiter la question du facteur de confusion.

Dans le second, Mantel et Haenszel présentent une nouvelle méthode, à laquelle ils ont laissé leur nom, qui est une méthode d'ajustement que l'on peut effectuer *a posteriori*, qui, en scindant les variables en classes (par exemple en tranches d'âge de 10 ans, selon Mantel et Haenszel [27]), et en estimant le risque relatif ou l'odds ratio par exemple par tranche d'âge, permet d'obtenir une mesure plus précise, un risque ajusté, pour l'ensemble des strates en ajoutant les différentes mesures par strates. Cette méthode constitue encore aujourd'hui un outil essentiel pour éliminer les facteurs de confusion dans une étude épidémiologique. Il faut aussi noter que c'est dans cet article qu'apparait la formulation mathématique, sans le mot, de l'*odds ratio* sous la forme R=AD/BC. En d'autres termes, l'année 1959 marque un progrès méthodologique majeur dans l'histoire de l'épidémiologie avec la formulation conceptuelle et mathématique de ces deux estimateurs. Enfin, Cornfield, comme Mantel et Haenszel considère de façon explicite que les études cas-témoin constituent une source de preuves aussi valide que les preuves expérimentales, contrairement aux thèses de Berkson, de Fisher ou plus tard de Sackett et Feinstein. Cornfield dit ainsi : «Il n'existe pas de catégories telles que des preuves de première classe et des preuves de seconde classe. (…) Distinguer entre, d'un côté, des associations statistiques et, de l'autre, des relations qui sont établies par l'expérimentation, sans aucune référence à des variables alternatives qui seraient présentes dans un cas mais pas dans l'autre ne

nous semble pas être de la bonne statistique, ni de la bonne science, ni de la bonne philosophie – bien que cela puisse constituer une bonne diversion » [28]. Mantel et Haenszel considèrent quant à eux que « la méthode rétrospective peut être considérée, selon la théorie statistique établie, comme la méthode d'étude de prédilection » [27].

Pour conclure sur ce point, en une dizaine d'années, les études cas-témoins ont acquis un statut scientifique important au sein de l'épidémiologie, malgré leurs défauts, qui sont connus des épidémiologistes. Les années 1960 et 1970 voient ces types d'études, du fait de leur faible coût et de leur rapidité de mise en œuvre, se multiplier [20] et surtout être utilisées pour justifier des politiques de santé publique. Inévitablement, des erreurs et des controverses se produisent (par exemple la controverse autour de la réserpine et du cancer du sein [29]) non seulement au sein de la communauté des épidémiologistes, entre les épidémiologistes et les autres scientifiques, mais aussi vis-à-vis du public en général, notamment en raison des « implications politiques profondes consécutives aux résultats de ces études » [14]. C'est dans cette « atmosphère intellectuelle de scepticisme » [14] qu'a lieu ce symposium sur les études cas-témoins.

2.3. De l'épidémiologie populationnelle à l'épidémiologie clinique, et retour

Le dernier point qui nous intéresse est que cette conférence oppose deux catégories ou plutôt deux générations d'épidémiologistes dont les parcours, passés et futurs, comme les objectifs poursuivis, diffèrent voire s'opposent : d'un côté, les représentants de l'épidémiologie populationnelle, comme J. Cornfield, A.M. Lilienfeld, P.E. Sartwell ; de l'autre, les fondateurs de l'épidémiologie clinique, notamment A.R. Feinstein et D.L. Sackett. L'épidémiologie clinique, c'est-à-dire une épidémiologie tournée non vers des populations mais vers les patients individuels, qui se préoccupe moins de santé publique que de médecine clinique – comme le souligne Sackett dans la définition qu'il en donne en 1969 [9] – commence en effet à s'institutionnaliser dans les années 1970-1980 [30] [10], ce qui ne manque pas de susciter des conflits : le mot de Lilienfeld qui qualifie, dans la discussion finale, ce symposium de «*Bermuda Summit Peace Conference*» [13] atteste de la violence de la controverse. La perspective des épidémiologistes cliniques sur ces études diffère largement de celle des épidémiologistes classique : Feinstein, comme Sackett, à qui il faut ajouter l'épidémiologiste A. Cochrane [31], s'attachent essentiellement à rechercher dans la

littérature médicale les études valides dans la perspective d'une médecine par essence individuelle.

Plus précisément, la ligne de fracture entre ces deux camps porte en réalité sur la validité des études cas-témoins et donc sur la capacité des études épidémiologiques observationnelles à établir une relation de causalité. Or, notre hypothèse est précisément que, alors même qu'apparait au cours de cette conférence la première définition du concept de biais par Sackett, cette définition ne correspond pas ou plus à ce que les tenants de l'épidémiologie populationnelle appellent un biais. En ce sens, les critiques qu'ils adressent à la méthodologie des études cas-témoins sont à la fois datées et injustifiées, datées car ces critiques ont déjà été faites dans les années 1950 et 1960 à l'occasion de la controverse sur le tabagisme et le cancer du poumon, et injustifiées car les épidémiologistes de l'époque (qui sont pour certains d'entre eux présents à la conférence) ont déjà répondu à ces critiques et les ont pour l'essentiel réfutées. Quelles sont donc ces critiques ?

Pour Feinstein, le problème de l'étude cas-témoins réside dans le fait que sa méthodologie n'est pas aussi rigoureuse que celle de l'expérimentation (une critique déjà faite par J. Berkson, discutée et réfutée par Hill [32] et Cornfield [28]) : «En épidémiologie, néanmoins, le genre particulier d'investigation de la relation de cause à effet que l'on appelle «l'étude cas-témoins» n'est pas conduite comme une expérimentation et sa performance aussi bien que son interprétation ne sont pas évaluées selon des standards scientifiques analogues [à la physique, la chimie et la biologie] » [33]. Sackett dit sensiblement la même chose, mais concentre sa critique sur les biais des études cas-témoins : «Les méthodologistes ont pendant trop longtemps ignoré la responsabilité qui leur incombe de mesurer l'occurrence et l'amplitude des biais, comme le montrent les trente années qui se sont écoulées entre la description du biais du taux d'admission [le biais de Berkson] et sa première démonstration empirique. Nous avons été justement critiqués pour cela, et nous devons nous mettre au travail» [1].

Cette référence au biais de Berkson, que Sackett a, avec d'autres collègues, démontré empiriquement en 1978 [34], n'est pas anodine : en effet, J. Berkson [24] [25] [35] a été, avec R.A. Fisher [23] [36], un des principaux critiques de la relation causale entre le tabagisme et le cancer du poumon. Cette controverse a été l'occasion d'une sévère critique de la méthodologie des études cas-témoins mais aussi des études de cohorte. Pour résumer cette critique, deux points ont été soulevés, le premier questionnant

l'association statistique entre tabagisme et cancer du poumon, le second la relation de causalité entre tabagisme et cancer du poumon :

(1) Le premier argument consiste à dire que l'association statistique observée est fausse ou fallacieuse (« *spurious* ») : en d'autres termes, cette association n'aurait pas de «signification biologique», et ne serait en fait qu'un «phénomène statistique artificiel», dû «à l'interaction de divers «biais» subtils et compliqués» [25]. Autrement dit, si l'on veut prouver une association entre deux variables, il faut procéder à une expérimentation, et des études observationnelles ne suffisent pas.

(2) Le deuxième argument consiste à dire que même si l'association était valide, elle ne signifierait pas qu'il y ait un lien de causalité entre tabagisme et cancer du poumon. L'argument, qui est aussi défendu par Fisher (dans une version plus génétique) ou par J. Yerushalmy [37], renvoie à ce qu'on peut appeler l'hypothèse du marqueur : la substance de cet argument est que le fait de fumer et le fait de développer un cancer auraient un facteur commun ou une cause commune, d'ordre physiologique, génétique ou comportemental, qui ferait que ces personnes auraient plus tendance à la fois à fumer et à développer un cancer du poumon. En d'autres termes, comme le souligne A. Fagot-Largeault, le tabac serait un « marqueur » du cancer et non un facteur[38].

Ici, seul l'argument (1) nous intéresse : la critique de Berkson porte en effet sur la méthodologie de l'étude [39], notamment sur le fait que, dans les études observationnelles et contrairement à l'étude expérimentale, la randomisation est impossible, d'où la référence aux biais. Or, sans randomisation, pour Berkson comme pour Fisher, il est impossible d'effectuer une « inférence statistique valide » [24]. C'est cette relation entre randomisation et biais qu'il convient à présent d'examiner.

3. Randomisation et biais :

3.1. Fisher, Hill et la randomisation

La notion de randomisation a été théorisée par Fisher dans les années 1920 [40] et apparait comme un « ingrédient essentiel à son approche du plan et de l'analyse des expérimentations»[41]. Ainsi, pour lui, le « but de la randomisation » est de « garantir la validité du test de signification» [42], au sens où cette randomisation neutralise les biais, c'est-à-dire les erreurs systématiques. Ainsi, dans le cadre d'une expérimentation agronomique sur la betterave fourragère qui vise à déterminer quel est le meilleur traitement parmi les cinq testés, en prenant comme critère le poids de la racine de betterave, Fisher montre qu'un arrangement systématique des parcelles, par

opposition à un arrangement au hasard, va certes «diminuer les erreurs réelles de l'expérimentation, mais au prix de l'augmentation de l'estimation de l'erreur » [42]. Si donc l'expérimentateur choisit un «système préarrangé», il est fort probable que «nos parcelles aient (…) des traits communs avec la variation systématique de fertilité, et alors notre test de signification est complètement vicié» [40]. La randomisation contribue ainsi à garantir la validité du test statistique de signification, test qui est fondé sur l'estimation de l'erreur garantie par la réplication des traitements. En d'autres termes, l'absence de biais devient chez Fisher la condition de possibilité de la validité du test statistique : « pour les tests de signification, un biais, s'il est prouvé, serait fatal »[43].

Néanmoins, en ce qui concerne l'introduction de la randomisation en épidémiologie, c'est bien à Hill qu'on la doit [41]. La fonction opératoire du concept de randomisation n'est pourtant pas identique à celle qu'elle a chez Fisher. En effet, pour Hill, tel qu'il l'exprime dans sa « Philosophie de l'essai clinique », publiée en 1953 [44], la randomisation vise essentiellement à « cacher la procédure d'allocation » des patients dans les différents groupes aux médecins, ce qui « garantit trois choses », résumées ainsi par J-P. Amann [45] : « 1) que les préférences personnelles des médecins chargés d'inclure les malades dans l'essai n'interviennent pas dans la constitution des groupes de traitement ; 2) que leurs efforts pour surmonter ces préférences n'introduisent pas de nouveaux biais ; et 3) que des critiques extérieurs ne puissent mettre en doute le caractère complètement aléatoire des affectations de patients dans les groupes ».

3.2. Le biais comme menace à l'objectivité de l'étude

Ainsi, si la fonction du concept de randomisation est différente chez Fisher et Hill, la fonction et le sens du concept de biais sont eux aussi différents : là où Fisher parle de biais statistique ou objectif, en relation avec la notion d'erreur systématique (par opposition à la notion d'erreur aléatoire), Hill semble plutôt se référer à la notion triviale, psychologique et subjective de la notion de biais, en relation avec le sens de « préjugé » qu'il a dans la langue anglaise. S'agit-il ici d'une réelle opposition de sens ou bien d'une simple stratégie rhétorique ? Les historiens de la médecine et de l'épidémiologie privilégient en général la seconde hypothèse. Ainsi, Harry Marks souligne qu'«à chaque fois qu'ils expliquaient la nécessité d'un changement de procédure — randomisation, essai en double aveugle, évaluation objective des résultats — les réformateurs invoquaient le mot de « biais». (…). Les chercheurs

étaient régulièrement mis en garde contre les dangers de biais «inconscients» pouvant invalider leurs études.» [46]. Cette hypothèse est cohérente avec la justification que donne Hill du choix qu'il a fait de ne pas utiliser le terme « randomisation » : « j'ai délibérément laissé de côté le mot « randomisation » (…) à cette époque-là, car j'essayais de persuader les médecins d'entrer dans les essais contrôlés de la manière la plus simple, et j'aurais pu les faire fuir » [6]. De fait, la principale différence entre les expérimentations de Fisher, qui portent essentiellement sur l'agriculture, et celles effectuées dans le cadre des essais cliniques randomisés, est que ces derniers incluent des êtres humains, ce qui soulève des difficultés d'ordre pratique et éthique supplémentaires. En effet, il serait compliqué et surtout immoral de pratiquer une expérimentation sur des êtres humains où par exemple on forcerait, dans un groupe, des sujets à fumer, et dans l'autre, des sujets à ne pas fumer, afin de déterminer les effets du tabagisme sur la santé.

Le problème se complique lorsqu'il s'agit des enquêtes étiologiques d'observation, rétrospective ou prospective. En effet, dans ces cas-là, la randomisation est tout simplement impossible : il est donc logique que le concept de biais ne puisse signifier la même chose dans ce contexte. Pourtant c'est bien la question de la validité de l'étude qui est posée, et spécifiquement la validité non pas de la relation de causalité, mais de l'association statistique elle-même, ce qui renvoie à la question du plan de l'étude. Ce glissement de la question de la causalité à celle de l'association est d'ailleurs visible dans les articles de Doll et Hill : alors que leur article de 1950 [21], qui porte sur une étude rétrospective, conclut que « le tabagisme est un facteur important dans l'étiologie du cancer du poumon », leur article de 1954, qui porte quant à lui sur leur étude prospective, conclut plus prudemment que «l'association entre tabagisme et cancer du poumon est réelle et non due à un biais»[47]. Cette attention aux biais est néanmoins présente dès les premières enquêtes, et même avant, notamment dans l'ouvrage de Hill de 1937 [48] : au moment où il s'agit pour les statisticiens de convaincre les médecins, que « les théories mathématiques n'impressionnent pas » [49], d'adopter une nouvelle méthodologie fondée sur les statistiques, il est préférable que cette méthodologie soit bien fondée.

Ainsi Doll et Hill examinent-ils dans leurs articles un certain nombre de biais, entendus comme une explication alternative à l'association entre tabagisme et cancer du poumon : sont examinés le biais de sélection, qui renvoie aussi bien à la représentativité de l'échantillon qu'à la comparabilité entre le groupe des cas et des témoins ; le biais de mémoire, ou encore le biais de l'intervieweur. Leur stratégie

argumentative repose sur le fait qu'ils examinent systématiquement les alternatives, pour mieux les invalider, et par là concluent que seule l'hypothèse du tabagisme est la bonne, ou du moins, qu'elle est la meilleure : «La possibilité que ces résultats soient dus au choix d'un groupe témoin inapproprié, au choix de patients atteints d'affections respiratoires surévaluant leur consommation de tabac ou à un biais dû à l'enquêtrice a été envisagée. Les raisons qui nous ont permis d'exclure toutes ces éventualités sont exposées, et on conclut de l'étude que le tabagisme est un facteur important dans l'étiologie du cancer du poumon. » [21].

3.3. Une nouvelle tripartition : cause, hasard, biais.

D. Mainland est sans aucun doute l'auteur qui a le plus travaillé sur ce concept de biais. Dans un article publié en 1955 apparait, à notre connaissance, la première définition explicite du concept de biais au sein de l'épidémiologie : les biais y sont définis comme «les choses qui modifient un échantillon de telle sorte qu'il en devient différent de ce qu'il est censé être » [50]. Cette définition, bien que vague en apparence, précise en réalité la pensée de Mainland, qui avait en 1953 défini le biais comme un «quelque chose» (dans son article, une compétition entre les taux de létalité de différentes maladies) qui vient «masquer une réelle association ou bien créer une association fallacieuse» [51]. En 1958, enfin, il reprend sa définition de 1955, et la clarifie : « le biais est tout ce qui rend un échantillon différent de ce qu'il est censé être – c'est une erreur d'étiquetage [*a mislabeling]* » [52]. Mainland, dans la continuité de Berkson et de Fisher, dont il était proche [53] [54], finit par assigner sa place et sa fonction au concept de biais dans la nouvelle économie conceptuelle qui est celle des enquêtes épidémiologiques observationnelles, par opposition aux expérimentations randomisées, en généralisant le concept de biais : « dans une enquête [d'observation] nous n'avons aucune protection contre des biais cachés. Dès lors, notre interprétation des résultats ne peut pas prendre la simple forme « ou bien-ou bien ». Nous devons considérer trois causes possibles des différences observées [entre deux échantillons] : (1) le hasard, (2) le facteur testé, et (3) des biais cachés, non dus au hasard » [55]. Ainsi, lorsqu'une étude cas-témoins ou de cohorte montre une association entre deux variables à travers une différence observée par exemple en termes de consommation de tabac et de prévalence de cancer du poumon, seules trois explications à cette différence sont possibles : le hasard, c'est-à-dire que l'association est purement fortuite – et le test statistique a pour fonction d'éliminer cette hypothèse en spécifiant le risque de se tromper ; la cause, c'est-à-dire que le facteur testé, ici le tabagisme, est un

facteur ou la cause de la différence de prévalence du cancer du poumon entre le groupe des cas et le groupe de témoins ; ou bien le biais, c'est-à-dire toute explication alternative à cette différence observée. Le problème est précisément qu'en l'absence de randomisation il est impossible de savoir laquelle de ces trois explications est correcte.

3.4. La critique de Hill par Mainland : biais subjectifs et biais objectifs

L'opposition entre l'approche du biais de Mainland et celle de Hill est ici plus claire : il ne s'agit pas simplement des biais psychologiques des médecins ou des intervieweurs, donc du sens trivial du mot biais que lui donne Hill, mais bien du biais comme catégorie générale qui recouvre toutes les explications alternatives à l'explication en termes de hasard ou de cause. Mainland critique d'ailleurs « la notion populaire de biais en tant que phénomène psychologique » qui « persiste chez certains enquêteurs médicaux » [52]. Ainsi, pour lui, par exemple, la « procédure des enveloppes scellées », qui vise à garantir l'allocation au hasard des patients dans différents groupes de traitement n'a pas comme chez Hill pour fonction essentielle de cacher la procédure aux médecins : « Ce type de procédure (…) est apparemment souvent considérée comme un dispositif pour éviter les biais psychologiques, conscients ou inconscients. Cette procédure joue un rôle beaucoup plus fondamental, qui n'a rien à voir avec les enveloppes scellées » [56], et qui vise précisément à permettre la randomisation et par là à effectuer une inférence statistique valide. Ainsi, même dans les expérimentations « il y a une masse immergée [Mainland vient de comparer la causalité en médecine à un « iceberg »] de facteurs potentiels de biais ; mais par la randomisation nous en contrôlons les effets par le hasard » [55].

Dans un article consacré à la controverse sur le rôle causal du tabagisme dans le cancer du poumon, M. Parascandola distingue deux manières de considérer la nature de la preuve apportée par les études épidémiologiques : « les opinions différentes sur les preuves reflétaient deux modèles différents de la recherche étiologique – l'expérimentation contrôlée comme le test crucial et objectif de l'hypothèse causale, par opposition au jugement inférentiel fondé sur un corpus divers de preuves » [57]. Parascandola considère d'ailleurs Berkson, Fisher et Mainland comme des partisans du test objectif de l'hypothèse causale via l'expérimentation contrôlée, tandis que Hill, mais aussi « Hammond, Cornfield et Lilienfeld » seraient plutôt partisans du jugement inférentiel fondé sur un ensemble de preuves diverses. Nous pensons que cette opposition, qui porte sur la question de la causalité, peut tout aussi bien s'appliquer à

celle de l'association, qui la précède d'un point de vue logique : dans la mesure où il est difficile, voire impossible, de recourir à la randomisation dans les études observationnelles, il est par conséquent impossible de décider si l'association est fallacieuse (« *spurious* ») ou réelle, sauf à examiner, comme Hill et Doll le font par exemple dans leur article de 1954, les explications alternatives pour aboutir à la meilleure explication possible.

4. Le concept de biais et l'épistémologie de l'épidémiologie

4.1. L'épidémiologie comme méthode

L'épidémiologie moderne, soit la période qui va de 1965 à aujourd'hui, selon la périodisation de A. Morabia [58], se caractérise essentiellement par son attention apportée à la méthode. En 2007, Rothman souligne ainsi qu'«aujourd'hui, nos défis ont l'air plus centrés sur la méthodologie ou la biologie, car nous concentrons toute notre attention à résoudre des biais insolubles ou à expliquer des résultats contradictoires » [59] ; tandis que A. Morabia se demande si l'épidémiologie n'est pas devenue obsédée par la méthode, oubliant ses objectifs de santé publique. A. Broadbent va plus loin et considère que l'épidémiologie se résume à une méthode ou à un ensemble de méthodes, au sens où «l'expertise d'un épidémiologiste est méthodologique» [60]. De fait, il peut paraître quelque peu paradoxal que Feinstein ou Sackett se plaignent de l'absence de standards méthodologiques et finalement de l'absence de scientificité de l'épidémiologie : si les critiques des années 1950 émises par Berkson, Fisher ou d'autres sont légitimes, car la méthodologie de l'épidémiologie en est encore à ses balbutiements, ce n'est plus le cas à la fin des années 1970. Comment interpréter alors ces critiques assez virulentes ?

En réalité, le procès qui est fait à l'étude cas-témoins, et aux études épidémiologiques observationnelles en général, ne peut se comprendre pleinement qu'en se référant aux intentions des auteurs. Plus précisément, les critiques de Feinstein ou de Sackett ne peuvent se comprendre que si l'on adopte leurs perspectives de recherche : en réalité, il ne s'agit pas pour eux de faire de l'épidémiologie une science mais plutôt de faire de la médecine clinique une science. En d'autres termes, il s'agit surtout pour eux d'éviter que les médecins commettent des erreurs, comme le souligne Cochrane [31], et c'est pourquoi ils se concentrent en général sur les questions de thérapeutique. Ainsi, dès 1963, Feinstein publie un article sur les « principes fondamentaux de la science clinique » [61], et son objectif est développer ce qu'il appellera, dans les années 1980, la « clinimétrie » [62]. Le projet de Feinstein est donc explicite dès le départ : il s'agit de

faire de la médecine clinique une science, au même titre que la médecine faite au laboratoire. Plus précisément, pour Feinstein, les études épidémiologiques bien conduites doivent partir des problèmes cliniques pratiques. Autrement dit, ni les populations ni les politiques de santé publique ne l'intéressent : la perspective est donc complètement différente par rapport aux épidémiologistes populationnels. Les questions et les méthodes de recherche sont donc elles aussi nécessairement différentes.

4.2. Une théorie du biais : Edmond Murphy et la logique de la médecine

De même, Murphy, dont la définition du biais a directement et explicitement inspiré celle de Sackett (Sackett fait d'ailleurs un compte-rendu élogieux de l'ouvrage de Murphy [63]), entend fonder une nouvelle « logique de la médecine », c'est-à-dire une théorie de l'inférence, car il considère que la médecine est en « crise » et qu'elle est insuffisamment « académique », c'est-à-dire insuffisamment scientifique. Son propos est donc beaucoup plus large que la question des études épidémiologiques, comme il l'avoue en citant A.B. Hill dans son introduction, pour souligner que le propos de Hill est «moins général que le sien» au sens où il porte essentiellement sur le problème d'effectuer «des inférences à partir des données épidémiologiques», et non d'effectuer des inférences à partir de toutes les données, y compris les données épidémiologiques. La nouveauté principale du traitement fait par Murphy du concept de biais consiste ainsi en deux points cruciaux : d'abord, il réunit et fusionne différentes significations disponibles à l'époque du mot biais, de la plus triviale (son sens psychologique de préjugé, au sens où «l'opinion d'un homme est considérée comme biaisée si elle ne représente pas fidèlement les faits») à la plus scientifique (il souligne que la différence entre le sens commun et l'usage du mot «biais» dans «le contexte épistémologique», où il est employé «de façon analogue» se situe en fait dans l'absence de «connotations morales» [3]), et balaie le spectre de ses nuances sémantiques pour aboutir à une épuration de ce qui fait l'essence même du concept de biais. Ensuite, il définit le biais comme un « processus » à travers lequel « l'appréhension et l'évaluation de la vérité scientifique pourraient être déformées », ce qui lui permet de distinguer différents biais en fonction des différentes étapes de la recherche, au nombre de six (Sackett, lui, en dénombre sept [1]) : le plan d'expérience, l'observation, l'estimation, le test des hypothèses, l'interprétation et le compte-rendu. Plus intéressant encore, Murphy montre non seulement que des biais peuvent surgir à chaque étape de la recherche scientifique, mais aussi qu'un « cercle vicieux du biais »

est susceptible de se produire et de modifier systématiquement et de façon erronée les connaissances scientifiques : après avoir théorisé le biais de publication, c'est-à-dire le fait de ne publier que les résultats positifs, qui est selon lui encore « plus marqué dans les revues prestigieuses qui circulent le plus », Murphy montre que «tous ces facteurs conspirent à déformer la croyance générale à propos de ce qui a été démontré dans un cas particulier et, dans la mesure où une telle croyance constitue le fondement de conviction préalables utilisées pour juger des futures publications, un cercle presque clos du biais peut se perpétuer» [3]. Sackett reprend cette idée dans son article : à la fin de la septième et dernière étape de la recherche (le biais de publication), il indique ainsi entre crochets : «*and back to (1)*» [1], c'est-à-dire retour à la première étape qui est celle où le chercheur s'informe de l'état du domaine ou du sujet.

4.3. Du concept médical au concept épidémiologique de biais

Ainsi, il semble assez évident que la question des biais chez Feinstein, Murphy ou Sackett, dépasse le simple cadre des études épidémiologiques et la question de la validité de l'association statistique ou de l'inférence causale : les études épidémiologiques apparaissent comme une source de données parmi d'autres, à côté des données expérimentales obtenues en laboratoire, des essais cliniques randomisés ou bien encore des données cliniques ; toutes ces données étant au service d'un soin bien-fondé aux malades individuels. La fonction opératoire du concept de biais est donc ici éminemment critique : il s'agit de mettre en garde les médecins contre toutes les formes d'erreurs qui sont susceptibles de vicier le raisonnement du médecin et par conséquent de lui faire commettre des erreurs dans sa pratique clinique. Il s'agit donc de cultiver une certaine forme de scepticisme, et le médecin-scientifique décrit par Murphy, au même titre que l'épidémiologiste clinique décrit par Sackett, doivent exercer un doute méthodique et permanent : à l'examen clinique des patients doit désormais s'ajouter l'examen critique des articles scientifiques (l'avant-dernier chapitre de l'ouvrage de Murphy est ainsi consacré à la lecture critique d'articles). C'est le sens aussi de la notion d'«évaluation critique » (« *critical appraisal* ») dont Sackett commence à enseigner les principes en 1978 [64]. Cette « évaluation critique » est ensuite renommée, en 1990, « *Evidence-based medicine* » [65] : un de ses principes fondamentaux est de hiérarchiser les preuves en fonction des types d'étude et de leur susceptibilité aux biais. Ainsi, moins une étude est affectée par des biais et plus elle est probante.

Le concept de biais tel qu'il est défini par Sackett, dans la continuité de Murphy, remplit donc une fonction opératoire différente et beaucoup plus large que celle que lui assignent les épidémiologistes. En effet, depuis les années 1980, les épidémiologistes opposent traditionnellement le concept de biais à celui de validité des études, et plus spécifiquement à celui de validité interne : Kleinbaum, Kupper et Morgenstern distinguent ainsi deux composantes de l'exactitude («*accuracy*»), ou plutôt deux sources d'inexactitude. L'erreur aléatoire pose un problème de précision («*precision*») alors que l'erreur systématique pose un problème de validité [15]. La validité se décompose elle-même en validité interne et validité externe : «La *validité interne* concerne la validité des inférences qui portent sur la population-cible, en utilisant des informations tirées de la population étudiée. (…). Le terme *validité externe*, quant à lui, concerne les inférences faites à propos de la population externe, donc au-delà de l'intérêt restreint de l'étude», ou en d'autres termes la généralisabilité des conclusions de l'étude. Dans les termes utilisés par les épidémiologistes des années 1950, la validité interne renvoie ainsi à la question de l'association statistique, tandis que la validité externe renvoie à celle de la causalité. En effet, si les résultats d'une étude sont généralisables à toute la population, cela signifie qu'ils sont valides non seulement pour l'échantillon de l'étude, mais aussi pour « la population au-delà de l'intérêt restreint de l'étude » [15]. Dès lors, il est possible d'effectuer un jugement de causalité, à condition bien sûr de tenir compte des résultats d'autres enquêtes ainsi que de « la connaissance théorique du processus de la maladie et des facteurs associés, et de considérations biologiques »[15]. Cette distinction entre validité interne et externe, que l'on retrouve chez Rothman et Greenland, qui considèrent que «le but général d'une étude épidémiologique consiste habituellement dans l'exactitude de l'estimation »[66], est aujourd'hui fondamentale en épidémiologie, au même titre que celle entre l'erreur aléatoire et l'erreur systématique, c'est-à-dire le biais. Si, comme le dit Morabia, « il n'y a pas de théorie du biais en 1945 » [58], ce n'est pas une théorie du biais qui est « établie » mais bien deux théories, ou plutôt deux fonctions opératoires du concept de biais : un concept épidémiologique, dont la fonction consiste à déterminer les menaces à la validité interne d'une étude, et un concept médical, dont la fonction consiste à rappeler aux médecins les vertus du scepticisme et du criticisme dans la pratique médicale.

5. Conclusion

Au terme de ce parcours historique, une question reste en suspens : pourquoi la définition du biais donnée par Sackett est-elle devenue la définition de référence en épidémiologie, alors qu'au même moment, ou presque, le biais est défini par les épidémiologistes comme une menace à la validité interne d'une étude ?

La réponse à cette question n'est pas simple et inclut sans doute des facteurs économiques et sociaux – liés notamment à l'institutionnalisation de l'épidémiologie comme discipline scientifique ; qu'historiques – liés à la multiplication des études épidémiologiques entre 1950 et 1979. A cela s'ajoutent des facteurs proprement épistémologiques : en à peine trente ans, l'épidémiologie a ainsi acquis un « statut de discipline scientifique grâce à une standardisation et une mathématisation accrue de ses méthodes », mathématisation parfois critiquée aujourd'hui [67,68] ; et a grandement contribué à rendre la médecine plus scientifique en permettant « d'identifier certaines causes de maladie », « d'évaluer la performance des méthodes diagnostiques, pronostiques et thérapeutiques » ou encore de faire régresser « de nombreuses maladies infectieuses » ainsi que « l'incidence de plusieurs maladies chroniques »[69].

Ainsi, alors même que les études épidémiologiques se sont multipliées depuis 1950, il est logique que les médecins, dans leur pratique clinique, puissent bénéficier des résultats de ces études et soient en même temps attentifs à la méthodologie de ces études ; d'où le concept médical de biais, en relation avec une hiérarchisation des études en fonction de leur niveau de preuves, ce qui inclut notamment leur susceptibilité aux biais. Ce privilège accordé par les tenants de l'« Evidence-based medicine » aux essais cliniques randomisés, comme seul type d'études valides, est d'ailleurs critiquable car il devrait logiquement conduire à l'impossibilité de conduire des études observationnelles, y compris celles dont le design est robuste : dès 1959, Cornfield [28], comme Hill d'ailleurs [32], avaient pourtant déjà défendu l'idée qu'il n'y avait pas deux catégories de preuves selon qu'elles soient issues d'une étude observationnelle ou d'une étude expérimentale. Ainsi, dans la perspective de l'épidémiologie clinique, l'épidémiologie est un moyen en vue d'une fin, au sens où c'est la médecine clinique qui est centrale. A l'inverse, les épidémiologistes ont continué de travailler sur la méthodologie de leur discipline, et ils étaient d'autant plus justifiés à le faire qu'on a longtemps accusé l'épidémiologie de ne pas être une science à part entière. Se focaliser sur la méthode a donc constitué un des seuls moyens pour l'épidémiologie de légitimer son statut de science : l'épidémiologie est donc ici une fin

en soi, et non un simple moyen, ce qui explique d'ailleurs que certains épidémiologistes lui reprochent d'avoir perdu de vue son objectif de santé publique.

Ce sont sans doute Rothman et Greenland qui expriment le mieux la pratique et l'objectif de l'épidémiologie, essentiellement tournés vers la méthodologie des études : «De la même manière qu'on ne peut utiliser aucun critère causal pour établir la validité d'une inférence, il n'y aucun critère utilisable pour établir la validité d'une donnée ou d'une preuve. (…) Mais bien qu'il n'y ait aucun critère absolu pour affirmer la validité d'une preuve scientifique, il reste néanmoins possible d'affirmer la validité d'une étude. Ce qui est requis est plus que l'application d'une liste de critères. Au contraire, il faut pratiquer une critique méticuleuse dans le but d'obtenir une évaluation quantifiée de toutes les erreurs qui affectent l'étude» [4]. Le vrai problème de toute étude épidémiologique n'est pas donc plus de savoir s'il y des erreurs ou des biais, mais de savoir combien, c'est-à-dire de les quantifier : «Le véritable problème est de quantifier l'erreur» [4].

Dès lors, à partir du moment où l'étude épidémiologique consiste, selon Rothman et Greenland, en un exercice de « mesure d'un effet causal», le concept de biais retrouve son sens originel d'erreur systématique de mesure, par opposition à l'erreur aléatoire, tel qu'il apparait chez les fondateurs de la théorie des erreurs: en effet, si les erreurs aléatoires sont éliminables aussi bien par le calcul différentiel (au sens où il existe des limites maximales d'erreur) que par le calcul des probabilités (au sens où les erreurs s'annulent ou se compensent), les erreurs systématiques, elles, restent constantes, y compris lorsque la taille de l'échantillon augmente, et s'accumulent au lieu de s'annuler. Si, au XVIIIe siècle, ce type d'erreurs apparait comme la « gangue métaphysique » de laquelle on extrait le « minerai de l'erreur fortuite » [70], c'est-à-dire comme un élément inanalysable ou incalculable, ce n'est plus le cas à la fin du XXe siècle, et ce n'est pas le moindre des tours de force de l'épidémiologie que d'avoir permis la définition, la classification et finalement la quantification des biais.

Pour conclure, une réflexion sur le passé, le présent et le futur de l'épidémiologie n'est sans doute pas inutile dans le contexte pandémique actuel. Si l'épistémologie de l'épidémiologie a été quelque peu « troublée » [71] lors de la première vague de la COVID-19, il semble aussi que la période soit propice à une réflexion sur la nature mais aussi sur les buts de l'épidémiologie : alors que l'épidémiologie clinique, tout comme l'épidémiologie des facteurs de risque, se concentrent essentiellement sur l'individu, et tandis que l'épidémiologie populationnelle se concentre essentiellement sur ses méthodes, notamment mathématiques, la pandémie de COVID-19 a montré

qu'il était sans doute temps pour l'épidémiologie de (re)devenir sociale : R. Horton considère ainsi que la COVID-19 n'est pas une pandémie mais une « syndémie » [72] qui révèle les interactions profondes entre le social et le biologique, et qui à la fois reflète les inégalités économiques et sociales et les aggrave du fait de politiques publiques inadaptées. En 1996, M. Susser and E. Susser [73] appelaient à changer de paradigme dominant en épidémiologie et à adopter le paradigme éco-épidémiologique afin de soutenir une épidémiologie orientée vers la santé publique. Il faut donc espérer que cette crise sanitaire puisse permettre cette révolution paradigmatique en épidémiologie.

Déclaration de liens d'intérêts L'auteur déclare ne pas avoir de liens d'intérêts.

Remerciements Je remercie les professeurs Joël Coste et Alain Leplège pour leur travail de relecture critique et constructive effectué sur une première version de cet article. Cet article est le fruit d'un travail de thèse effectué au sein du Laboratoire SPHERE, UMR 7219, Université Paris Diderot : je remercie les membres de cette unité, notamment la directrice Mme Sabine Rommevaux-Tani, pour leur soutien tout au long de ce travail.

Références

[1] Sackett DL. Bias in analytic research. J Chronic Dis 1979;32:51–63. https://doi.org/10.1016/0021-9681(79)90012-2.

[2] Last JM, International Epidemiological Association, editors. A dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001.

[3] Murphy EA. The logic of medicine. Baltimore: Johns Hopkins University Press; 1976.

[4] Rothman KJ, Greenland S. Causation and Causal Inference in Epidemiology. Am J Public Health 2005;95:S144–50. https://doi.org/10.2105/AJPH.2004.059204.

[5] Vineis P. History of bias. Soz- PräventivmedizinSocial Prev Med 2002:156–61.

[6] Hill SAB. Memories of the British streptomycin trial in tuberculosis. Control Clin Trials 1990;11:77–9. https://doi.org/10.1016/0197-2456(90)90001-I.

[7] Streptomycin Treatment of Pulmonary Tuberculosis. A Medical Research Council Investigation. Br Med J 1948;2:769–82.

[8] Sackett DL. Interview with David Sackett (2014-2015) n.d.:104.

[9] Sackett DL. Clinical Epidemiology. Am J Epidemiol 1969;89:125–8.

[10] Giroux É. De l'épidémiologie de santé publique à l'épidémiologie clinique. Quelques réflexions sur la relation entre épidémiologie et clinique (1920-1980). Bull D'histoire D'épistémologie Sci Vie 2012;Volume 19:21. https://doi.org/10.3917/bhesv.191.0021.

[11] Rockhill B. The privatization of risk. Am J Public Health 2001;91:365–8. https://doi.org/10.2105/ajph.91.3.365.

[12] Canguilhem G. Le concept et la vie. Rev Philos Louvain 1966;64:193–223. https://doi.org/10.3406/phlou.1966.5347.

[13] Ibrahim MA, editor. The Case-Control Study: Consensus and Controversy. Pergamon Press; 1979.

[14] Ibrahim MA, Spitzer WO. The case control study: The problem and the prospect. J Chronic Dis 1979;32:139–44. https://doi.org/10.1016/0021-9681(79)90026-2.

[15] Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and methods. New York: Wiley; 1982.

[16] Rothman KJ. Modern epidemiology. 1st ed. Boston: Little, Brown; 1986.

[17] Paneth N, Susser E, Susser M. Origins and early development of the case-control study: part 1, Early evolution. Soz PräventivmedizinSocial Prev Med 2002;47:282–8. https://doi.org/10.1007/PL00012638.

[18] Paneth N, Susser E, Susser M. Origins and early development of the case-control study: part 2, The case-control study from Lane-Claypon to 1950. Soz PräventivmedizinSocial Prev Med 2002;47:359–65. https://doi.org/10.1007/s000380200003.

[19] Lilienfeld AM, Lilienfeld DE. A century of case-control studies: Progress? J Chronic Dis 1979;32:5–13. https://doi.org/10.1016/0021-9681(79)90004-3.

[20] Cole P. The evolving case-control study. J Chronic Dis 1979;32:15–27. https://doi.org/10.1016/0021-9681(79)90006-7.

[21] Doll R, Hill AB. Smoking and carcinoma of the lung: preliminary report. Br Med J 1950;2:739–48. https://doi.org/10.1136/bmj.2.4682.739.

[22] Omran AR. The epidemiologic transition. A theory of the epidemiology of population change. Milbank Mem Fund Q 1971;49:509–38.

[23] Fisher RA. Smoking.The cancer controversy: some attempts to assess the evidence. Edinburgh, Scotland: Oliver and Boyd; 1959.

[24] Berkson J. The statistical study of association between smoking and lung cancer. Proc Staff Meet Mayo Clin 1955;30:319–48.

[25] Berkson J. Smoking and Lung Cancer: Some Observations on Two Recent Reports. J Am Stat Assoc 1958;53:28. https://doi.org/10.2307/2282563.

[26] Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions*. Int J Epidemiol 2009;38:1175–91. https://doi.org/10.1093/ije/dyp289.

[27] Mantel N, Haenszel W. Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease 1959;22:30.

[28] Rubin E. Questions and Answers. Am Stat 1954;8:19. https://doi.org/10.2307/2681548.

[29] Horwitz RI. Exclusion Bias and the False Relationship of Reserpine and Breast Cancer. Arch Intern Med 1985;145:1873. https://doi.org/10.1001/archinte.1985.00360100139023.

[30] Zimerman A L. Evidence-Based Medicine: A Short History of a Modern Medical Movement. AMA J Ethics 2013;15:71–6. https://doi.org/10.1001/virtualmentor.2013.15.1.mhst1-1301.

[31] Cochrane AL. Effectiveness and efficiency: random reflections on health services. London: Nuffield Provincial Hospitals Trust; 1972.

[32] Hill AB. Observation and Experiment. N Engl J Med 1953;248:995–1001. https://doi.org/10.1056/NEJM195306112482401.

[33] Feinstein AR. Methodologic problems and standards in case-control research. J Chronic Dis 1979;32:35–41. https://doi.org/10.1016/0021-9681(79)90009-2.

[34] Roberts RS, Spitzer WO, Delmore T, Sackett DL. An empirical demonstration of Berkson's bias. J Chronic Dis 1978;31:119–28. https://doi.org/10.1016/0021-9681(78)90097-8.

[35] Berkson J. Smoking and Lung Cancer. Am Stat 1963;17:15–22.

[36] Stolley PD. When Genius Errs: R. A. Fisher and the Lung Cancer Controversy. Am J Epidemiol 1991;133:416–25. https://doi.org/10.1093/oxfordjournals.aje.a115904.

[37] Yerushalmy J, Palmer CE. On the methodology of investigations of etiologic factors in chronic diseases. J Chronic Dis 1959;10:27–40. https://doi.org/10.1016/0021-9681(59)90015-3.

[38] Valleron A-J. L'Épidémiologie humaine: conditions de son développement en France, et rôle des mathématiques. Les Ulis; [Paris: EDP Science ; Académie des sciences; 2006.

[39] Brault N. Le biais de Berkson, ou l'histoire d'un quiproquo épistémologique. Bull D'histoire D'épistémologie Sci Vie 2020;Volume 27:31–49. https://doi.org/10.3917/bhesv.271.0031.

[40] Fisher RA. Statistical Methods for Research Workers. 1ère. Edinburgh, Scotland: Oliver and Boyd; 1925.

[41] Armitage P. Fisher, Bradford Hill, and randomization. Int J Epidemiol 2003;32:925–8. https://doi.org/10.1093/ije/dyg286.

[42] Fisher RA. The Design of Experiments. Oliver & Boyd; 1935.

[43] Neyman J, Iwaszkiewicz K, Kolodziejczyk S. Statistical Problems in Agricultural Experimentation. Suppl J R Stat Soc 1935;2:107. https://doi.org/10.2307/2983637.

[44] Hill AB. The Philosophy of the Clinical Trial. Stat. Methods Clin. Prev. Med., Edinburgh and London: E. & S. Livingston; 1962, p. 3–14.

[45] Amann J-P. Chapitre 5. La philosophie de l'essai clinique selon Austin Bradford Hill. L'émergence Médecine Sci., Editions Matériologiques; 2012, p. 113. https://doi.org/10.3917/edmat.fagot.2012.01.0113.

[46] Marks HM. Confiance et méfiance dans le marché : les statistiques et la recherche clinique (1945-1960). Sci Soc Santé 2000;18:9–27. https://doi.org/10.3406/sosan.2000.1503.

[47] Doll R, Hill AB. The mortality of doctors in relation to their smoking habits: a preliminary report. Br Med J 1954;1:1451–5. https://doi.org/10.1136/bmj.1.4877.1451.

[48] Hill AB. Principles of Medical Statistics. Londres: The Lancet Ltd; n.d.

[49] Kaptchuk TJ. Intentional Ignorance: A History of Blind Assessment and Placebo Controls in Medicine. Bull Hist Med 1998;72:389–433. https://doi.org/10.1353/bhm.1998.0159.

[50] Mainland D. Use of Case Records in the Study of Therapy and Other Features in Chronic Disease: I. Planning the Survey. Ann Rheum Dis 1955;14:337–52. https://doi.org/10.1136/ard.14.4.337.

[51] Mainland D. The risk of fallacious conclusions from autopsy data on the incidence of diseases with applications to heart disease. Am Heart J 1953;45:644–54. https://doi.org/10.1016/0002-8703(53)90306-8.

[52] Mainland D. Notes on the planning and evaluation of research, with examples from cardiovascular investigations. Part I. Am Heart J 1958;55:644–55. https://doi.org/10.1016/0002-8703(58)90002-4.

[53] Mainland D. The Rise of Experimental Statistics and the Problems of a Medical Statistician. Yale J Biol Med 1954;27:1–10.

[54] Altman D. Donald Mainland: anatomist, educator, thinker, medical statistician, trialist, rheumatologist. J R Soc Med 2020;113:28–38. https://doi.org/10.1177/0141076819896274.

[55] Mainland D. Notes on the planning and evaluation of research, with examples from cardiovascular investigations. Part II. Am Heart J 1958;55:824–37. https://doi.org/10.1016/0002-8703(58)90099-1.

[56] Mainland D. Notes on the planning and evaluation of research, with examples from cardiovascular investigations. Part III. Am Heart J 1958;55:838–50.

[57] Parascandola M. Two approaches to etiology: the debate over smoking and lung cancer in the 1950s. Endeavour 2004;28:81–6. https://doi.org/10.1016/j.endeavour.2004.02.003.

[58] Morabia A. Has Epidemiology Become Infatuated With Methods? A Historical Perspective on the Place of Methods During the Classical (1945–1965) Phase of Epidemiology. Annu Rev Public Health 2015;36:69–88. https://doi.org/10.1146/annurev-publhealth-031914-122403.

[59] Rothman KJ. Commentary: Epidemiology still ascendant. Int J Epidemiol 2007;36:710–1. https://doi.org/10.1093/ije/dym151.

[60] Broadbent A. Philosophy of epidemiology. Palgrave Macmillan UK; 2013.

[61] Feinstein AR. The basic elements of clinical science. J Chronic Dis 1963;16:1125–33. https://doi.org/10.1016/0021-9681(63)90048-1.

[62] Feinstein AR. An Additional Basic Science for Clinical Medicine: IV. The Development of Clinimetrics. Ann Intern Med 1983;99:843. https://doi.org/10.7326/0003-4819-99-6-843.

[63] Sackett DL. The Logic of Medicine. By Edmond A. Murphy. Baltimore: Johns Hopkins Press, 1976. Pp. 353. $16.50. J Med Philos 1977;2:71–6. https://doi.org/10.1093/jmp/2.1.71.

[64] Haynes RB, Goodman SN. An interview with David Sackett, 2014–2015. Clin Trials J Soc Clin Trials 2015;12:540–51. https://doi.org/10.1177/1740774515597895.

[65] Guyatt G. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. JAMA 1992;268:2420. https://doi.org/10.1001/jama.1992.03490170092032.

[66] Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. Philadelphia Baltimore New York: Wolters Kluwer Health, Lippincott Williams & Wilkins; 2008.

[67] Broadbent A. Causation and prediction in epidemiology: A guide to the "Methodological Revolution." Stud Hist Philos Sci Part C Stud Hist Philos Biol Biomed Sci 2015;54:72–80. https://doi.org/10.1016/j.shpsc.2015.06.004.

[68] Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. Int J Epidemiol 2016;45:1776–86. https://doi.org/10.1093/ije/dyv341.

[69] Leplège A, Bizouarn P, Coste J. De Galton à Rothman: les grands textes de l'épidémiologie au XXe siècle. Paris: Hermann; 2011.

[70] Armatte M. Statut de la Dispersion : de l'erreur à la variabilité. Journl Électronique D'Histoire Probab Stat Electron J Hist Probab Stat 2010;6:20.

[71] Coste J, Bizouarn P, Leplège A. L'épistémologie troublée de la première vague de recherche sur la Covid-19. Rev DÉpidémiologie Santé Publique 2020. https://doi.org/10.1016/j.respe.2020.09.001.

[72] Horton R. Offline: COVID-19 is not a pandemic. The Lancet 2020;396:874. https://doi.org/10.1016/S0140-6736(20)32000-6.

[73] Susser M, Susser E. Choosing a Future for Epidemiology: II. From Black Box to Chinese Boxes and Eco-Epidemiology. Am J Public Health 1996;86:674–7.