

Les métamorphoses de la preuve dans le champ psychiatrique

L'exemple des études cliniques sur les psychotropes

Revue d'anthropologie des connaissances, 7, n°3, 2013

Éric Trombert

Abstract

Este artículo es una contribución al estudio de las transformaciones contemporáneas en el campo de la psiquiatría a partir de un análisis de la investigación en psicofarmacología basado en el análisis de la estructura y de los estilos de publicaciones en este campo. En él se destacan tres etapas principales: (1) la edad de oro de los grandes descubrimientos (litio, clorpromazina, imipramina), documentada por una descripción clínica en cada caso, (2) el incremento del ensayo aleatorio controlado y del foco puesto en una metodología para evitar la incertidumbre y los sesgos clínicos, (3) los últimos años, marcados por la competencia entre las moléculas, el desarrollo de métodos meta-analíticos y el aumento de la opacidad de los artículos para el lector clínico. Esta contribución se inscribe en el contexto de una epistemología histórica de las prácticas clínicas.

Introduction

Les protagonistes de ce qu'on a appelé l' « Evidence-Based Medicine » (EBM) poursuivent leur objectif de développement d'une médecine toujours davantage fondée sur les « preuves », soit l'idée selon laquelle les décisions médicales devraient par principe se baser sur les résultats d'études cliniques et empiriques ayant démontré leur « efficacité » (Evidence-Based Medicine Working Group, 1992). Dans ce cadre, les études empiriques et cliniques accessibles aux médecins – et qui constituent dès lors les preuves supposées fonder leurs actions – sont donc essentiellement formées de travaux publiés dans les revues médicales dites « scientifiques » et dont la valeur est estimée

aujourd'hui selon trois critères principaux : la plus haute valeur probante revenant aux essais randomisés contrôlés et, de plus en plus

fréquemment, aux méta-analyses de tels essais², le nombre de cas étudiés et le facteur d'impact des journaux dans lesquels ces travaux sont publiés. En prenant un peu de recul historique, comme nous allons le montrer, les articles publiés évoluent, à la fois sur le plan méthodologique et sur le plan formel. Ce qui constitue aujourd'hui une « preuve », au sens que lui donne le paradigme de l'EBM, est très différent de ce que l'on pouvait trouver dans les revues scientifiques il y a une cinquantaine d'années. Nous pouvons même faire l'hypothèse que des études cliniques dont l'impact a été majeur seraient aujourd'hui considérées comme peu scientifiques et ne seraient pas publiées. Nous nous proposons ici d'étudier trois temps marquants de l'évolution du domaine dont nous sommes le plus proche en tant que psychiatre clinicien : la psychopharmacologie.

Le premier temps, de la fin des années 1940 à la fin des années 1960, est celui de la découverte des premiers psychotropes considérés comme efficaces pour traiter les pathologies psychiatriques majeures, rencontrées dans les asiles : la manie, la schizophrénie et la mélancolie, à savoir les psychoses classiques³. Nous montrerons en particulier que les études sur le lithium, la chlorpromazine et l'impramine sont méthodologiquement fondées sur l'observation clinique de série de cas, une caractéristique dont découle une forme narrative et argumentative basée sur la description qualitative de cas singuliers.

Le deuxième temps, des années 1970 à la fin du siècle, est celui de l'apparition de l'usage de la statistique comme outil de preuve de l'efficacité d'une molécule, désormais évaluée au cours d'essais dits randomisés et contrôlés, dont l'application en médecine s'étend progressivement à partir des années 1940. Il s'agit là d'un phénomène complexe aux enjeux multiples (avec des dimensions épistémologiques, politiques, industrielles et professionnelles, liste non exhaustive ; Gaudillière, 2006 ; Marks, 1997). En psychiatrie, les essais randomisés contrôlés ont impliqué en parallèle le développement des échelles psychométriques de sévérité, ainsi qu'une standardisation progressive des procédures expérimentales, notamment une approche systématique du diagnostic dont on retrouve les principes au cœur du projet du DSM-III⁴, dont la parution a nettement orienté la recherche clinique psychiatrique à partir des années 1980. Les articles issus de cette démarche apparaissent d'une grande simplicité formelle. Ils peuvent souvent se résumer à un tableau chiffré, censé démontrer l'efficacité mesurée d'une molécule par rapport à un placebo.

Le troisième temps couvre ces dix dernières années ; elles sont marquées par une baisse tendancielle de l'innovation psychopharmacologique et se caractérisent par une mise en concurrence de molécules voisines que les auteurs des études publiées s'efforcent de départager en termes d'efficacité. La démarche implique alors une complexité croissante des protocoles des études empiriques et une sophistication poussée des méthodes statistiques permettant d'évaluer des efficacités comparées. Il en résulte des publications qui nous paraissent plus techniques et finalement assez opaques pour les cliniciens.

Comme nous allons le voir au cours de cet article, ce troisième temps est problématique à plusieurs égards pour le clinicien se préoccupant des fondements épistémologiques de la construction des preuves censées orienter son action.

Temps 1. L'âge d'or – *decennium mirabilis*

Entre 1949 et 1957, sont découvertes coup sur coup les trois classes majeures de psychotropes : les stabilisateurs de l'humeur (le lithium en 1949), les antipsychotiques (la chlorpromazine en 1954) et les antidépresseurs (l'imipramine en 1957).

L'étude originale rapportant l'efficacité des sels de lithium pour traiter « l'excitation psychotique » est publiée dans le *Medical Journal of Australia*, une revue plutôt modeste sur le plan international. L'auteur, John F. Cade, rapporte l'évolution de dix cas traités par des sels de lithium, dont il avait constaté l'effet sédatif chez le cochon d'Inde, animal qu'il utilisait pour étudier l'excitation psychotique. Les cas y sont décrits un à un, de façon détaillée mais non systématique. Il est utile de le citer pour rendre compte de la forme et de la structure de l'argument :

W.-B., a male, aged fifty-one years, who had been in a state of chronic manic excitement for five years, restless, dirty, destructive, mischievous and interfering, had long been regarded as the most troublesome patient in the ward. His response was highly gratifying. From the start of treatment on March 29, 1948, with lithium citrate he steadily settled down and in three weeks was enjoying the unaccustomed surroundings of the convalescent ward. As he had been ill so long and confined to a "chronic ward", he found normal surroundings and liberty of movement strange at first. Owing to this, as well as housing difficulties and the necessity of determining a satisfactory maintenance dose, he was kept under observation for two months. He remained perfectly well and left hospital on July 9, 1948, on indefinite leave with instructions to take a maintenance dose of lithium carbonate, five grains twice a day. The carbonate had been substituted for the citrate as he had become intolerant of the latter, complaining of severe nausea. He was soon back working happily at

his old job. However, he became lackadaisical about his medicine and finally ceased taking it. His relatives reported that he had not had any for at least six weeks prior to readmission on January 30, 1949, and was becoming steadily more irritable and erratic. He ceased work just before Christmas. On readmission to hospital he was once again started on lithium carbonate, ten grains three times a day, and in a fortnight had been settled down to normal. The dose of carbonate was then reduced to five grains three times a day, and in a further two weeks to five grains twice a day. He is now (February 28, 1949) ready to return to home and work. (Cade, 1949, p. 350)

Comme en témoigne cet extrait, la description est colorée, elle assume une part d'appréciation subjective et inclut aussi certains jugements de valeur. Il faut également relever qu'elle suit l'évolution temporelle du cas, donnant ainsi l'impression que l'auteur connaît bien son patient et le suit de près dans le cadre d'une clinique régulière et rapprochée. Les autres cas de l'étude sont décrits de façon similaire. Leur présentation ne suit pas de canevas systématique : la description des symptômes, les aspects biographiques, l'évolution clinique, les aspects posologiques sont souvent mentionnés dans le désordre et de manière plus ou moins prolixe, renforçant l'impression d'une interpénétration forte entre la démarche expérimentale et le suivi clinique rapproché. Le diagnostic est toujours lié à un état maniaque mais il est formulé de façon aussi variable, sans explication quant à la signification du diagnostic ou la façon dont il a été posé⁵. Pour les dix cas rapportés, l'efficacité des sels de lithium paraît impressionnante.

Trois ans plus tard, les psychiatres français Jean Delay, Pierre Deniker et Jean-Marie Harl présentent à la Société Médico-Psychologique leur expérience avec une molécule mise au point par Henri Laborit, dans le cadre de ses recherches anesthésiologiques : la chlorpromazine (Delay, Deniker et Harl, 1952a), dont ils décrivent un mois plus tard l'effet sur huit cas « d'excitation et d'agitation ».

Accès d'excitation atypique récidivante : S... Robert, 21 ans, étudiant, est transféré de l'hôpital Beaujon le 21-11-1951 pour sub-excitation atypique avec procursivité nocturne. On note une tendance au rationalisme morbide et quelques troubles du cours de la pensée.

Le malade est suivi depuis 1947 à Marseille, pour accès maniaques rapprochés et de plus en plus atypiques. Des rémissions transitoires ont été obtenues par des séries d'électro-chocs, une cure cardiazolique et insulinique.

Dans la famille, on note un cousin maniaco-dépressif. L'examen est négatif. Dans le service, le malade est en proie à une grande agitation avec insomnie. Devant l'atypicité des troubles, on fait une cure de 32 comas insuliniques, qui amène une rémission relative en mars, le malade restant subexcité. Trois électro-narcoses calment le malade pendant quinze jours, mais l'agitation confuse, les violences, les déclamations, la logorrhée euphorique contraignent de nouveau à

maintenir le malade insomnique [sic].

Le traitement par le 4560 RP permet dès les premiers jours le retour au calme, coupé de quelques phases de logorrhée. Après 21 jours de traitement, tous les symptômes d'excitation, mais aussi les signes de dissociation mentale et le rationalisme morbide ont disparu. Le malade peut sortir au début de juin, paraissant tout à fait normal.

Le médicament a donc permis d'obtenir une rémission meilleure chez ce malade que par l'insuline, les chocs ou les électro-narcoses. La manie, très atypique, nous faisait craindre une schizophrénie. (Delay, Deniker et Harl, 1952b, p. 272)

La description apparaît plus complexe sur le plan psychopathologique que dans l'étude de John F. Cade et certains termes peuvent même être assez obscurs pour le lecteur d'aujourd'hui (procurtivité, par exemple). La présentation est aussi plus systématique : symptômes, antécédents, réponse au traitement. Les diagnostics dans la série sont plus variés, et là non plus ils ne sont pas explicités ou discutés⁶, même si, d'un point de vue clinique, on perçoit le lien entre les symptômes décrits et le diagnostic retenu. Sur le plan méthodologique et formel, les démarches des psychiatres français et de John F. Cade sont très similaires sur le plan clinique. La lecture des résultats rapportés est là aussi impressionnante.

Dans les mêmes années, Heinz E. Lehmann et Gorman E. Hanrahan publient dans les *Archives of Neurology and Psychiatry* le premier article en anglais rapportant leur expérience sur soixante et onze cas traités par chlorpromazine. Leur approche est pourtant un peu différente des deux précédentes :

Patients receiving the drug become lethargic. Manic patients often will not object to bed rest, and patients who present management problems become tractable. Assaultive and interfering behavior ceases almost entirely. The patients under treatment display a lack of spontaneous interest in the environment, yet are easily accessible and respond as a rule immediately and relevantly to questions even if awakened from sleep. [...] Some patients dislike the treatment and complain of their drowsiness and weakness. Some state that they feel "washed out", as after an exhausting illness, a complaint which is indeed in keeping with their appearance. Others, however, appreciate the peculiar effect of the drug which allows them to relax from ideational pressure. One manic patient stated after her recovery that with the drug she soon lost the feeling: "I had to live my whole life in one day". Another patient suffering from long standing anxiety described the effect of small doses of chlorpromazine on her feelings: "It was like a chairman taking control of a meeting where everybody previously had been shouting at once". (Lehmann et Hanrahan, 1954, p. 230-231)

H.-E. Lehmann et G.-E. Hanrahan font référence à un ensemble de

patients désormais trop important pour qu'ils soient décrits cas par cas et l'évolution temporelle de chacun ne peut plus être détaillée. Les auteurs proposent plutôt une synthèse générale de leurs observations, qu'ils illustrent d'exemples et de citations de ce que les patients disent eux-mêmes de leur expérience. L'approche est non systématique, mais les auteurs construisent cependant une synthèse structurée de leurs résultats, sous la forme d'un tableau rapportant l'évolution des cas – en quatre degrés non définis : « unchanged », « improved », « much improved », « recovered » – et selon huit catégories diagnostiques, elles aussi non définies dans l'article : « schizophrenia, schizoaffective », « manic, acute », « psychoneurotic », etc. Malgré cette approche plus systématique, les auteurs semblent bien connaître leurs patients, dont ils peuvent citer les propos.

En 1957, la très modeste *Schweizerische medizinische Wochenschrift* publie le premier article rapportant l'action de l'imipramine dans le traitement des états dépressifs. L'auteur, Roland Kuhn, travaille à la clinique de Münsertlingen dans le canton de Turgovie, en Suisse. Il cherche lui aussi un traitement pour ses patients schizophrènes et demande à la société Geigysi si elle dispose d'une molécule expérimentale apparentée à la chlorpromazine que le clinicien pourrait éventuellement expérimenter. L'entreprise accepte et met à sa disposition l'imipramine. Roland Kuhn remarque bientôt que, si cette molécule a assez peu d'efficacité antipsychotique, elle semble améliorer les symptômes dépressifs.

L'article de Kuhn se présente lui aussi sous une forme synthétique. L'auteur ne précise pas le nombre de patients traités, mais il est très élevé⁷. Il ne lui est donc pas possible de présenter l'évolution de chaque cas. La méthode et la forme sont ainsi très similaires à l'article de Lehmann et Hanrahan. L'efficacité du traitement est décrite dans deux paragraphes qu'il convient de citer *in extenso* pour rendre compte de la manière dont l'auteur rapporte, structure et commente ses observations :

Les symptômes dépressifs visibles de l'extérieur s'améliorent sous G 22355 de manière remarquable. La mimique se relâche et retrouve ses capacités de modulation et d'expression. Les malades deviennent de manière générale plus vivants, le chuchotement dépressif se fait plus audible, les malades apparaissent plus bavards, la lamentation et les pleurs cessent. Lorsque, dans la dépression, on constatait une humeur d'insatisfaction plaintive ou d'hésitation, celle-ci disparaît, laissant apparaître un sujet plus sympathique, satisfait et ouvert. Les plaintes hypocondriaques ou neurasthéniques diminuent ou cessent totalement. Des malades qui ne voulaient pas sortir du lit le matin se lèvent plus tôt, avec les autres patients. Ils entrent à nouveau en relation avec les autres, commencent à converser, participent à la vie du service, reprennent une correspondance et se préoccupent à nouveau de leur famille. Ils reprennent spontanément un travail,

parviennent à certains succès, et leur lenteur laisse la place à un tempo de vie normal. Avec cette évolution, les malades sont rapidement appréciés dans l'unité et ils apparaissent plus équilibrés du point de vue de leur humeur et de leur comportement. À plusieurs reprises, des proches ont expliqué avec enthousiasme au médecin qu'ils n'avaient pas vu le malade aller aussi bien depuis longtemps.

La plupart du temps, mais pas toujours, les malades s'aperçoivent des changements eux-mêmes, ils les reconnaissent, en sont clairement très heureux et parlent de « cure miracle ». Les sentiments de pesanteur, de fatigue, d'asthénie, d'oppression, de tension interne, de rigidité ou d'inquiétude se relâchent. Les malades se sentent à nouveau libres, l'inhibition de la pensée et du comportement diminue, les pensées et les activités reviennent. Une humeur triste, opprimée, désespérée ou anxieuse devient neutre-insouciant ou discrètement enjouée, avec le sentiment de retrouver la santé et une force croissante ; les sentiments pathologiques de culpabilité, les idées de ruine et de péché disparaissent simplement ou perdent leur impact émotionnel, paraissent plus lointaines, le malade leur est plus indifférent. Une pulsion suicidaire disparaît parfois brusquement ! Si le sommeil était perturbé dans le cadre de la symptomatologie dépressive, il se normalise rapidement, sans faire recours à des hypnotiques, même dans les cas qui n'avaient pas montré de réponse aux somnifères usuels. Les « rêves pesants » que racontent parfois les dépressifs, rêves de sang, de corps morts, d'accidents horribles ou de tortures cruelles, souvent associés à une terrible angoisse, n'apparaissent plus sous traitement. L'aggravation matinale et les autres variations diurnes de l'état dépressif disparaissent rapidement. S'il y avait un manque d'appétit, il s'améliore également, de même parfois pour une constipation spécifiquement liée à l'état dépressif. (Kuhn, 1957, p. 1136)⁸

L'article de Kuhn est complexe. Il inclut des observations psychopathologiques, où il relève que la molécule est plus efficace dans les cas qu'il considère comme « endogènes »⁹, et des considérations épistémologiques, parfois assez poussées, ainsi que la description chimique de la molécule, et une discussion de ses liens avec la chlorpromazine et de ses effets secondaires. Il s'agit à nouveau d'une évocation synthétique, dans laquelle Kuhn s'intéresse aux symptômes qu'il considère comme les plus importants dans la dépression « endogène », à savoir l'aspect psychomoteur au sens large, l'élan vital, et les ruminations classiques des états mélancoliques¹⁰.

Ces quatre articles ont beaucoup en commun : une description clinique soigneuse mais peu systématique, sans procédure définie, et une approche descriptive, qui laisse une large part à l'appréciation subjective du clinicien et, dans les deux derniers cas, une place accordée aux patients eux-mêmes, cités directement ou indirectement. La démonstration de l'efficacité des molécules testées

repose ainsi sur ce type de description et ne fait appel à aucune procédure statistique. Elle se révèle largement favorable à ces nouvelles molécules dont l'impact sur les pathologies les plus sévères de l'asile psychiatrique est montré comme impressionnant. À les regarder avec les considérations en vigueur aujourd'hui avec la médecine des preuves, ces études ne répondent pas aux exigences contemporaines, non seulement sur le plan méthodologique, mais aussi sur le plan formel. En effet, dans le paradigme de l'EBM, le résultat qu'apporte l'article scientifique est considéré comme devant être accessible sans que l'auteur dudit article ne s'interpose entre le lecteur et la connaissance « objective » qu'il cherche à apporter, comme l'illustrent ces recommandations très normatives récemment publiées sur « l'écriture scientifique ». On peut y voir l'idéal d'une écriture désobjectivée¹¹ :

Une œuvre littéraire romanesque a, pour l'auteur, comme ambition de faire passer des émotions, des idées, des sentiments, parfois des connaissances au travers de personnages ou de situations plus ou moins inventés. [...] Les qualités d'écriture de [l'auteur] ayant un rôle prépondérant pour la bonne transmission de tous ces éléments.

La littérature scientifique est à l'opposé de tout cela. Aucune émotion ne doit exister dans un article scientifique. Un auteur scientifique ne regrette pas, il ne déplore pas ou ne s'attriste pas d'un résultat. Les mots « hélas », « heureusement » ou « par bonheur » n'ont pas leur place dans la littérature scientifique. Celle-ci ne doit apporter qu'une connaissance objective. Le lecteur ne doit pas se poser la question de savoir ce que l'auteur a bien pu vouloir dire [...]

James Farfor, un médecin écossais, a comparé la lecture scientifique à un aquarium : la vitre représente la forme et le contenu de l'aquarium, le fond. Si la vitre n'est pas claire le contenu est opaque. Si la vitre est parfaitement transparente, l'intérieur est bien visible. Si un article est mal rédigé son contenu même intéressant est difficile à saisir. Au contraire, correctement rédigé le contenu d'un article est plus facilement assimilable. (Cannac, Viargues et Dot, 2010, p. 4)

Pour Cannac et ses collègues, il y aurait sans doute « trop de littérature » chez Cade et ses collègues, pour que leurs arguments soient aujourd'hui considérés comme présentés de manière adéquate au style « scientifique » : des résultats décrits comme très probants, certes, mais peu de rigueur, peu de procédure explicite, pas de statistiques, et ce dans des revues bien modestes. Les publications scientifiques des décennies suivantes vont voir progressivement ces caractéristiques s'inverser.

Temps 2. Randomiser, mesurer, comparer

Les molécules découvertes durant les années 1950 vont s'imposer progressivement dans la clinique, mais non sans controverses. Le

lithium, en particulier, se heurte à un scepticisme durable, quant à son efficacité tant thérapeutique que prophylactique (Blackwell et Shepherd, 1968). De nombreux détracteurs américains et britanniques du lithium mettent en doute les résultats de Cade et de ceux qui l'ont suivi, notamment des psychiatres scandinaves. Ces débats s'inscrivent plus généralement dans le contexte de l'émergence de l'essai randomisé contrôlé comme seul dispositif permettant d'établir véritablement l'efficacité d'une molécule (Meldrum, 2000). Une conférence donnée à l'école de médecine de Harvard en 1952 par A. B. Hill¹² (1952) illustre bien la position des « réformateurs »¹³. Comme H. Marks (1997) l'a montré, ces exigences méthodologiques procèdent d'une triple méfiance à l'égard des cliniciens quant à (1) leurs théories sur la physiopathologie et le mode d'action des molécules ; (2) leur capacité à constituer des groupes de patients homogènes, sans biais de sélection¹⁴ ; (3) leur capacité à évaluer de façon fiable¹⁵ et objective¹⁶ l'action de la molécule. En raison de l'effet placebo, le compte rendu par les patients est considéré comme potentiellement biaisé en faveur du traitement et le patient doit donc lui aussi rester dans l'ignorance de ce qu'il reçoit (la procédure est dite « en double aveugle », *double-blind* en anglais : clinicien et patient ignorent si c'est la molécule testée ou la molécule contrôle que le patient reçoit)¹⁷. En d'autres termes, les études de Cade, Kuhn et les autres, fondées sur l'observation directe et minutieuse de patients sélectionnés par l'investigateur lui-même, sur des bases peu explicites et sans groupe contrôle illustrent tout ce qu'il ne faut pas faire aux yeux de des « réformateurs », dont les efforts visent à faire de la médecine en général et de la psychiatrie en particulier une pratique « scientifique », supposée pouvoir être libérée des croyances et idiosyncrasies des cliniciens, ou encore de l'influence des firmes pharmaceutiques. Comme nous l'avons mentionné dans l'introduction, le développement des essais randomisés contrôlés implique davantage que les questions méthodologiques qu'il semble simplement venir résoudre. Entre autres perspectives possibles, citons : l'émergence d'une nouvelle catégorie d'experts issus de l'épidémiologie et de la statistique, avec les luttes de pouvoir et de prestige au sein du monde médical en général et académique en particulier (Marks, 1997), et l'importance d'acteurs non médicaux comme les autorités de régulation, l'industrie pharmaceutique et divers groupes de pression de la société civile (Gaudillière, 2006). Il pose aussi de nombreux problèmes épistémologiques, notamment sur la question de ce qui vaut pour preuve. Pour notre propos, ce qui importe néanmoins est qu'il se soit imposé en médecine.

Les premiers essais incluant une forme de randomisation, utilisant un placebo et cherchant à documenter de façon plus méthodique l'évolution clinique des patients traités furent laborieux. Mogens Schou, un psychiatre danois qui s'intéresse beaucoup au lithium, publie en 1954 une étude portant sur trente-huit patients maniaques,

dont il évalue l'état clinique par une gradation de 1 à 3 des niveaux émotionnel et moteur (Schou, Juel-Nielsen, Stromgren et Voldby, 1954). Dans certains cas, le traitement est donné de façon ouverte, dans d'autres cas les patients reçoivent soit du lithium soit un placebo, sur une durée « d'habitude » de deux semaines. L'article comprend des schémas d'évolution pour quatorze patients considérés comme représentatifs, qui mettent en évidence de manière graphique l'effet du lithium, et l'absence d'effet du placebo, sur les symptômes. La durée d'observation et la procédure expérimentale sont encore très variables d'un patient à l'autre, mais cette étude adopte bien les prémisses des principes de l'essai randomisé contrôlé devenus aujourd'hui standards). Il implique un contrôle des « biais » attribués aux cliniciens, par la randomisation et leur ignorance du traitement réellement reçu par les patients, ainsi que des procédures explicites pour déterminer qui sont les patients inclus dans l'étude, mesurer de façon numérique l'évolution clinique et juger du résultat final par un test statistique sur les résultats chiffrés.

L'étude de Robert F. Prien et Jonathan O. Cole (1968) est l'une des premières études que nous avons trouvée qui adopte une stratégie explicite sur chacun de ces points, sauf le diagnostic. Ils mènent ce projet pour le compte du National Institute of Mental Health (l'autorité américaine pour la recherche en psychiatrie), dans sept hôpitaux psychiatriques publics, et l'étude sera publiée dans la principale revue de psychiatrie de l'époque. Elle vise avant tout à clarifier la question du dosage utile de la chlorpromazine dans le traitement de la schizophrénie. Les auteurs commencent leur article en posant sept questions, tout à fait spécifiques :

1. Is a high dose of 2,000 mg of chlorpromazine administered over a prolonged period of time more effective than the low dose of 300 mg typically used in the treatment of chronic schizophrenia?
2. Is high dose chlorpromazine more advantageous than treatment routinely administered to the chronic schizophrenic in the public mental hospital?
3. Are the two dose levels of chlorpromazine and the routine hospital treatment more effective than an inert placebo?
4. Do the treatments operate differentially on specific schizophrenic symptoms? For example, is high dose more effective than low dose in reducing delusional thinking?
5. Is there a type of patient who responds to one treatment and not to another?
6. What proportion of chronic patients show an exacerbation of symptoms following withdrawal of regular medication and initiation of placebo?
7. Are there differences between treatments in the nature and/or frequency of side effects? (Prién et Cole, 1968, p. 483)

Il vaut la peine de décrire par le détail le protocole adopté qui

deviendra bientôt un standard. L'étude est ambitieuse et porte sur huit cent trente-huit patients. L'article décrit soigneusement la procédure. L'échantillon remplit les critères suivants : diagnostic primaire de schizophrénie ; présence de symptômes actifs (dans au moins une de six catégories spécifiées : troubles des perceptions, pensée délirante, etc.) ; âge de 19 à 55 ans ; hospitalisation continue d'au moins deux ans, *i. e.* sans interruption dépassant huit semaines (critères d'inclusion) ; pas de signe de maladie cérébrale organique (y compris la lobotomie), ni de retard mental (QI en dessous de 70), ou de contre-indications médicales à l'utilisation de hautes doses (critères d'exclusion). Les patients ont ensuite été répartis au hasard (« randomized ») entre quatre groupes de traitement : « haute dose », « basse dose », « placebo », « traitement de routine ». Après huit semaines, les patients des trois premiers groupes changeaient de traitement (« switch »). L'évaluation est faite à l'aide d'échelles spécifiques ayant elles-mêmes fait l'objet de développements et de publications spécifiques. Les échelles sont remplies par les psychiatres (« Inpatient Multidimensional Psychiatric Scale », 81 items psychopathologiques), les infirmiers (« Nurse's Observation Scale for Inpatient Evaluation », 80 items décrivant le comportement du patient dans le service) et les assistants sociaux (« Discharge Readiness Inventory », 62 items évaluant dans quelle mesure un patient est prêt à quitter l'hôpital). Les psychiatres remplissaient également une évaluation globale de sévérité et d'amélioration cliniques (sur deux échelles de 1 à 7) et une évaluation des effets secondaires en 40 items. Les résultats sont présentés sous la forme de tableaux assez complexes, compte tenu du nombre de groupes et d'instruments et de sous-analyses en fonction de l'âge et de la durée d'hospitalisation. Les résultats thérapeutiques sont pourtant peu impressionnants sur l'ensemble du collectif : 10 % de patients avec une amélioration marquée dans le groupe « haute dose », contre 6 % dans le groupe « basse dose » et 3 % dans les deux derniers, par exemple). L'étude conclut à la supériorité de hautes doses de chlorpromazine.

L'article de Prien et Cole n'a ainsi pas grand-chose en commun avec les études des pionniers de la psychopharmacologie. Les auteurs ne connaissent pas les patients ; une procédure explicite doit garantir que le projet puisse se dérouler, dans des conditions relativement similaires, dans sept établissements différents ; le jugement final n'est prononcé ni par les cliniciens ni par les patients, mais par l'opération d'un calcul statistique. Les années qui suivent seront marquées par l'intégration de nombreuses méthodes d'évaluation psychométriques, puis de procédures diagnostiques qui se standardiseront elles aussi, fondées sur l'approche opérationnelle développée par le groupe de l'Université de Washington à Saint-Louis et reprise par le DSM-III (Schwartz et Wiggins, 1986).

Il faut toutefois relever que la plupart des essais randomisés contrôlés sont nettement plus simples que l'étude de Prien et Cole. Un exemple

parmi d'innombrables autres, l'étude de J. B. Cohn et C. Wilcox publiée en 1985 dans une importante revue de psychiatrie visant à comparer la fluoxétine (la molécule du Prozac® et de la Fluctine®) à l'imipramine et un placebo (Cohn et Wilcox, 1985). Les critères d'inclusion précisent que les patients ont été diagnostiqués « selon les critères du DSM-III », une formule désormais canonique dans son contenu¹⁸. Dans les années suivantes, l'étape diagnostique sera souvent encore plus explicite : les chercheurs sont invités à établir les diagnostics à l'aide d'« instruments » validés¹⁹, c'est-à-dire que l'on juge l'évaluation clinique, même référée aux critères DSM-III ou IV, insuffisamment rigoureuse pour être considérée comme scientifique. Comme dans l'étude de Prien et Cole, l'efficacité se mesure sur des échelles psychométriques spécifiques²⁰, largement utilisées et reconnues²¹. Sous cette forme idéale, l'essai randomisé contrôlé est un texte simple, illustré de deux ou trois figures ou tableaux eux aussi plutôt simples –, dont le contenu varie peu. À l'instar des sciences de la nature, selon un standard qui s'est largement généralisé, toutes ces études respectent un même canevas formel : résumé, introduction, méthode, résultat, discussion. Elles sont rapidement lues et aisées à comprendre.

Temps 3. La guerre des molécules, la guerre des compagnies

La fin des années 1980 est marquée, d'une part, par le triomphe assez inattendu du DSM-III²², qui prend une importance dans la clinique, l'enseignement et la recherche que les deux éditions précédentes n'avaient jamais eue ; ce succès touche aussi de nouveaux publics, au-delà du cercle des psychiatres spécialistes. On assiste, d'autre part et de manière concomitante, à l'arrivée sur le marché de la fluoxétine²³, un nouvel antidépresseur qui rencontre un succès stupéfiant sous le nom de Prozac®. À la croisée de ces deux phénomènes, la psychiatrie fait ainsi une entrée assez remarquable sur la scène publique. D'un côté, la littérature scientifique, dans laquelle on publie essentiellement des essais cliniques assez simples contre placebo ou contre des psychotropes plus anciens, cherche à démontrer l'efficacité des nouvelles molécules contre la dépression et la schizophrénie ; d'un autre, s'affirme une littérature plus « laïque » qui commente le succès du Prozac®²⁴ et vante volontiers les progrès supposés de la psychiatrie contemporaine (Schumer, 1989). Dans ce climat que l'on peut qualifier d'enthousiaste, le marché des psychotropes se développe de manière notoire²⁵. Les acteurs politiques, individuels et institutionnels, ne sont pas insensibles à ce climat : le président américain G. Bush déclare les années 1990 à 2000 « décennie du cerveau » et annonce un ambitieux programme de recherche neuroscientifique devant permettre de développer de nouvelles méthodes de traitement et de prévention pour les maladies neurodégénératives comme la maladie de Parkinson,

les traumatismes du système nerveux, les crises épileptiques et les troubles dépressifs.

Au terme de cette « décennie du cerveau », les cliniciens disposent de deux « nouvelles générations » de psychotropes : les inhibiteurs sélectifs de la recapture de la sérotonine (SSRIs) et de la noradrénaline (SNRIs), ainsi que les antipsychotiques atypiques et quelques « cousins pharmacologiques ». Des études cliniques ont fait la preuve contre placebo de l'efficacité de ces molécules, mais avec pour principal avantage comparatif, moins un gain important en termes d'efficacité qu'un avantage au niveau de leurs effets secondaires²⁶. Ces études n'ont cependant pas permis de comparer ces produits entre eux, qui paraissent d'une efficacité très similaire. Dans ce contexte, le National Institute of Mental Health finance quatre grandes études, ambitieuses et coûteuses, afin de déterminer les meilleures stratégies thérapeutiques parmi les nouvelles options, ce dans des conditions qui se rapprochent au mieux du « monde réel » (« real-world studies »). Trois pathologies psychiatriques majeures seront ainsi ciblées : le trouble bipolaire (étude STEP-BD), la dépression (études STAR*D et TADS) et la schizophrénie (étude CATIE) (National Institute of Mental Health, 2006).

Nous présentons ici à titre d'exemple l'étude CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness), menée pour « déterminer les effets à long terme et l'utilité des médications antipsychotiques chez les personnes avec une schizophrénie [...] qui pourraient bénéficier d'un changement de médication » (ClinicalTrials Feeds, 2006)²⁷. Son objectif est de déterminer la meilleure stratégie thérapeutique lorsqu'un antipsychotique se révèle inefficace ou mal toléré, et ceci dans des conditions se rapprochant de celles de la pratique clinique habituelle. CATIE est un projet complexe, sophistiqué sur le plan méthodologique : 1 493 sujets ont été inclus et ont, dès lors, reçu un antipsychotique, choisi au hasard parmi quatre nouveaux (« atypiques ») et un ancien (« classique ») ; les sujets ont ensuite été suivis sur 18 mois. Sur le plan de l'efficacité clinique des traitements, on peut considérer les résultats de cette étude comme extrêmement décevants : 74 % des sujets ont interrompu leur traitement dans un délai moyen de moins de quatre mois²⁸ (Lieberman, Stroup, McEvoy *et al.*, 2005). Mais surtout, cette étude n'est pas parvenue à démontrer de différences notables entre les divers produits, notamment entre les atypiques et le médicament classique, ce qui était son objectif principal. CATIE a ainsi été une surprise pour les cliniciens et une déception pour les entreprises pharmaceutiques commercialisant ces nouveaux antipsychotiques. Les résultats des études STAR*D pour la dépression et STEP-BD pour le trouble bipolaire²⁹ ont été elles aussi décevantes sur le plan de l'efficacité des molécules, et ne sont pas parvenues à les départager.

Les nombreux articles qui rendent compte de ces études apparaissent ainsi bien plus complexes que les simples essais randomisés contrôlés des décennies précédentes, tant au plan méthodologique qu'au niveau des résultats, sont désormais présentés dans de grands tableaux difficiles à lire, même pour les professionnels, et fatalement incomplets tant la quantité d'informations est importante, rendant aux auteurs l'exhaustivité impossible³⁰.

Si la simple lecture de ces résultats est peu aisée pour le clinicien, leur appréciation l'est plus encore, et ce d'autant plus qu'ils sont volontiers présentés sous un « jour flatteur ». Il est à cet égard intéressant de rapporter l'éditorial accompagnant la principale publication sur CATIE, parue dans une prestigieuse revue :

The value of CATIE is that it provides solid evidence to help clinicians and their patients make the difficult decisions needed to optimize the treatment of schizophrenia with the compounds currently available. (Freedman, 2005, p. 1288)

Cette présentation de CATIE est particulièrement problématique dans la mesure où, comme nous l'avons montré (Saraga et Stiefel, 2011), elle a précisément échoué à déterminer quelle était la meilleure des stratégies cliniques, en montrant au contraire qu'elles étaient toutes également assez médiocres.

Après cet échec relatif, apparaissent de nouvelles façons d'aborder cette même question de la détermination comparative de la meilleure molécule pour des pathologies données. Des procédures statistiques hautement sophistiquées permettent désormais de mettre les molécules concurrentes en situation de « compétition virtuelle », en intégrant dans le calcul toutes les études publiées les ayant comparées les unes aux autres en vue de produire pour résultat un classement d'efficacité. La prestigieuse revue *Lancet* a ainsi publié en 2009 une étude qui conclut à la supériorité de l'escitalopram et de la sertraline (deux SSRI, *i.e.* des « cousins » de la fluoxétine) sur leurs concurrents (Cipriani, Furukawa, Salanti *et al.*, 2009). Mais force est de constater que ce genre de conclusion reste problématique : l'avantage démontré est pour le moins marginal et il faut garder à l'esprit qu'il s'agit du résultat d'un calcul complexe rendu nécessaire par l'absence d'études cliniques concluantes sur la question. Cette procédure dite « multiple-treatments meta-analysis »³¹ pose de nombreux autres problèmes parmi lesquels des biais de publication, l'inclusion d'études sponsorisées, l'absence de prise en compte des doses, ou encore des résultats formulés comme un classement à la signification réelle douteuse puisque la procédure statistique implique que chaque molécule trouve son rang même avec un faible degré de certitude. De plus, les résultats figurent sur des tableaux devenus difficilement compréhensibles³² par un clinicien et la procédure statistique dépasse quant à elle de très loin les compétences ordinaires d'un non-

statisticien.

Le même groupe de recherche a publié en 2011 une étude similaire sur le traitement des épisodes maniaques dont l'un des résultats marquants était la mauvaise performance du lithium qui, en termes d'efficacité, en ressortait comme nettement inférieur à certains antipsychotiques (Cipriani, Barbui, Salanti *et al.*, 2011). Le niveau de compétence statistique nécessaire à la lecture critique de cette étude est tel qu'il faut un effort considérable et la consultation de la centaine de pages annexées sur Internet pour comprendre que la faible performance du lithium résulte en fait de deux études qui lui sont défavorables : une déjà ancienne contre l'halopéridol (7 patients dans chaque groupe) et une plus récente, contre l'halopéridol et la rispéridone (15 patients dans chaque groupe)³³. Ces deux études ne regroupent finalement que 59 patients. En tout, l'analyse de Cipriani *et al.* a inclus 68 essais randomisés contrôlés pour un total de plus de 16 000 patients ; 15 autres études comparant le lithium à un antipsychotique ou un autre stabilisateur de l'humeur étaient incluses dans l'analyse, et elles étaient toutes soit favorables au lithium soit neutres. Par exemple, la comparaison entre le lithium et le valproate, un antimaniaque tout à fait reconnu, était nettement favorable au lithium. Ces 59 patients devaient pourtant peser bien lourd d'un point de vue statistique, puisqu'au terme de la procédure, le lithium est jugé inférieur, en termes d'efficacité, à l'halopéridol, la rispéridone et l'olanzapine, trois antipsychotiques. Sur cette base, les auteurs concluent que les recommandations cliniques devraient être révisées, notamment en retirant au lithium sa place de traitement de première ligne de l'état maniaque. D'un point de vue clinique, on peut rester sceptique devant cette injonction à écarter de la pratique une molécule aussi éprouvée que le lithium, quelle que soit la sophistication statistique de l'étude.

Ces dix dernières années sont aussi marquées par une concurrence très rude entre les firmes pharmaceutiques pour un marché en expansion. Il apparaît clairement que de nombreuses études relèvent autant d'enjeux commerciaux que scientifiques ou cliniques, malgré leur sophistication technique. Car pour entrer sur le marché, une nouvelle molécule doit désormais faire la preuve de son efficacité face à ses « concurrents », preuve construite notamment par des moyens statistiques. Et en la matière il existe de nombreuses façons de favoriser un produit contre un autre, certaines plus subtiles que d'autres. La plus classique consiste à choisir une dose trop faible du comparateur, c'est-à-dire de la molécule adverse. Il arrive ainsi souvent que les grandes études complexes que nous avons décrites soient exploitées par les industriels pour mettre en avant leur produit en insistant sur un aspect spécifique qui lui a été favorable (même si ces études sont relativement irréprochables quant à leur indépendance par rapport aux firmes impliquées). En d'autres termes,

l'industrie fait également un usage situé et intéressé des résultats produits. Ce point vaut la peine d'être illustré. L'étude CATIE a montré une durée de traitement légèrement plus longue sous olanzapine, ce que son fabricant, Elli-Lilly, n'a pas manqué de souligner, mais aussi une prise de poids moindre sous quetiapine, ce que les représentants d'AstraZeneca, qui la produit, rappellent volontiers. Quant à l'étude de la revue *Lancet* sur les antidépresseurs, elle est toujours citée par le délégué de médical de Janssen-Cilag, qui insiste sur le bon résultat de l'escitalopram, sous patente, sans trop s'attarder sur le résultat équivalent de la sertraline, dont la patente est échue depuis plusieurs années déjà.

Les firmes pharmaceutiques doivent aussi lutter contre la concurrence des génériques, une fois la licence d'une molécule échue. Une manière de le faire est d'agir sur la forme galénique³⁴ pour conserver un marché acquis. Par exemple, la patente de la quetiapine (Seroquel®) est échue en mars 2012. Depuis plusieurs années, en réponse aux exigences de la Food and Drug Administration, AstraZeneca sponsorise de nombreuses études montrant l'efficacité d'une « forme retard » de la molécule, le Seroquel XR®, dans des indications qui sont déjà bien établies pour le Seroquel® standard (par exemple, Cutler, Datto, Nordenheim *et al.*, 2011). Le Seroquel XR® n'est pas plus efficace que le Seroquel® standard, ni que ses concurrents sur le marché des antipsychotiques ; il se prend simplement une fois par jour au lieu de deux ou trois fois pour le Seroquel® standard. C'est certes là un avantage potentiel pour les patients, mais il est bien modeste – alors que l'enjeu pour la firme AstraZeneca, qui pourra conserver sous patente pendant plusieurs années la même molécule sous une forme galénique un peu différente, est majeur. Pour cette dernière, il s'agit de faire passer un maximum de patients à la forme retard afin d'anticiper sur l'extinction de la patente de la forme standard, qui donnera lieu aussitôt à la production de versions génériques, faisant ainsi fondre le chiffre d'affaires important du Seroquel®. La stratégie n'est pas difficile à comprendre mais reste que, face aux études sur le Seroquel XR®, le clinicien qui est aussi le prescripteur ne saisira que difficilement l'aspect essentiellement commercial d'une démarche qui suit toutes les règles méthodologiques constitutives de la fabrication des preuves scientifiques dans la médecine d'aujourd'hui.

Terminons par un dernier point sur la question de l'implication de l'industrie dans la production de preuves. Au cours de ces dernières années, le rôle problématique du sponsoring est apparu de façon plus claire. Les universités ainsi que les organisations gouvernementales comme le National Institute of Mental Health aux États-Unis ou le Fonds National de la Recherche Scientifique en Suisse ont tendance à considérer que les études contre placebo relèvent plutôt de la responsabilité des firmes pharmaceutiques qui en exploitent les brevets. Le principal argument en la matière est que ces études

exigées pour l'homologation se rapprochent davantage d'une étape dans le développement d'un produit commercial que d'un effort scientifique à part entière. Or le sponsoring d'études cliniques explique probablement en partie le phénomène du biais de publication, bien démontré pour les études sur les antidépresseurs. En effet, les études publiées sont dans leur grande majorité favorables aux molécules testées, alors que c'est l'inverse pour les études non publiées (Turner, Matthews, Linardatos, Tell et Rosenthal, 2008). Notons toutefois que, depuis 2005, de nombreuses revues scientifiques refusent de publier des études n'ayant pas été enregistrées au préalable, ce qui devrait permettre de corriger progressivement ce biais (De Angelis, Drazen, Frizelle *et al.*, 2004).

Conclusion

La psychopharmacologie est une discipline relativement jeune, dont l'histoire s'inscrit évidemment dans celle, plus large, de son temps, au carrefour d'évolutions et de contraintes multiples. Ses méthodes ont évolué avec celles de la recherche médicale en général, passant d'une approche essentiellement clinique centrée sur l'observation méticuleuse des patients à des méthodes d'analyse donnant lieu à des études dans lesquelles les patients sont désormais observés de loin, et leurs caractéristiques cliniques transformées en nombres et analysées par des chercheurs sans lien direct avec eux. Ce développement est celui de l'Evidence-Based Medicine. Elle s'est imposée comme le modèle contemporain d'une médecine qui revendique sa « scientificité » et, partant, considère avec une certaine méfiance ce que peut penser, observer et déduire le clinicien impliqué dans la prise en charge singulière de son patient. La force de conviction du nombre est sans doute à situer dans un contexte sociopolitique plus large. En psychopharmacologie, comme dans les sondages d'opinion, les grands nombres (Desrosières, 1993) emportent la conviction dans les approches méta-analytiques qui consistent *in fine* à décider d'une vérité selon une forme de votation au cours de laquelle chaque étude voit son vote pondéré par le nombre de sujets qu'elle inclut et ce, sous le couvert d'une apparente lisibilité et le pouvoir de séduction qui lui est lié³⁵. L'Evidence-Based Medicine soulève encore de nombreux problèmes philosophiques (par exemple, les implications de la numérisation de l'humain qu'elle exige pour appliquer ses principes statistiques), épistémologiques (par exemple, sa prétention à constituer la seule démarche scientifique en médecine, dont atteste l'extension sémantique de « evidence » vers la « preuve », alors que c'est une notion en anglais plus proche de « indice ») et politiques (par exemple, proposer une démarche supposée scientifique et ainsi indemne d'enjeux politiques pour arbitrer les décisions cliniques – on voit ainsi se développer un mouvement nommé « Evidence-Based Policy » (Haynes, Service, Goldacre et Torgerson, 2012 ; Labrousse, 2010)). Notre propos est ici plus modeste, et vise seulement à discuter

ce que les transformations formelles des articles médicaux laissent voir des remaniements plus vastes qu'implique ce mouvement.

Le développement de l'Evidence-Based Medicine va de pair, en psychopharmacologie tout du moins – et peut-être de plus en plus largement en médecine – avec une stérilité assez frappante de la recherche : les essais randomisés contrôlés apportent aux cliniciens un certain nombre d'informations utiles, mais fort peu de nouveautés thérapeutiques. Certains auteurs ont d'ailleurs discuté de l'existence d'un lien de causalité entre ces deux constats. D. F. Klein (2008) explique cette stérilité par la perte de la « sérendipité », cette disponibilité à se laisser toucher par un hasard heureux qui caractériserait les chercheurs cliniciens vigilant et ouverts des années 1950-1960. Selon cet auteur, le phénomène serait lié au manque de temps, en particulier à la brièveté des séjours hospitaliers, mais aussi au modèle même de l'essai randomisé contrôlé dont l'approche par la moyenne s'oppose à un effort de compréhension de ce qui distingue entre eux les sujets répondant favorablement ou défavorablement à une molécule. Il propose d'ailleurs un autre modèle de recherche, plus proche des études de Cade, Kuhn, Delay, Deniker et Harl. D'autres critiques plus sophistiquées de l'approche randomisée contrôlée vont dans le même sens (par exemple et parmi bien d'autres, Jobe, Helgason et Roitberg, 2001 ; de Leon, 2012).

Un dernier point d'inquiétude est lié à cette opacité plus récente des études, dont les résultats sont souvent présentés de façon péremptoire et dont l'éventuelle finalité commerciale reste évidemment discrète. Le clinicien, pour conserver une posture critique, doit désormais être doté de solides compétences épistémologiques et statistiques, et l'on peut douter que ce soit fréquemment le cas. En se contentant de suivre les recommandations de pratique clinique et les conclusions sommaires des grandes études, il pratiquera une médecine fondée sur des preuves dont la valeur probante est souvent discutable, mais finalement peu discutée.

Notes

[1.](#) Pour une brève présentation de l'article et du dossier thématique dans lequel il s'inscrit, nous renvoyons le/la lecteur/lectrice à l'article introductif de Bovet, Kraus, Panese, Pidoux et Stücklin, « Les neurosciences à l'épreuve de la clinique et des sciences sociales. Regards croisés ».

[2.](#) Nous revenons sur ces termes par la suite.

[3.](#) Nous laissons ici de côté la question des « tranquillisants mineurs », mécloprobamate et benzodiazépines, dont la trajectoire nous paraît très différente, destinés plutôt à des patients non hospitalisés,

prescrits par des généralistes, et dont la contribution à la constitution de la psychopharmacologie, comme champ de compétence et comme promesse de développement de la psychiatrie, semble nettement moindre (cf. sur ce point, Herzberg, 2011).

4. 3^e édition en 1980 du Diagnostic and Statistical Manual of Mental Disorders publié par l'Association américaine de psychiatrie (APA). Le DSM-III définit des critères qui se veulent précis et relativement simples à appliquer pour chaque trouble mental, dans le but d'améliorer la concordance diagnostique (aussi nommée « fidélité inter-juge »).

5. Par exemple : « chronic mania », « recurrent mania », « manic-depressive insanity associated with alcoholism ».

6. « Accès maniaque suivi de confusion fébrile », « grande excitation confuso-délirante », « accès maniaque au cours d'une cyclothymie atypique », etc.

7. Un article paru un an plus tard parle de plus de 500 sujets traités par imipramine (Kuhn, 1958).

8. Nous traduisons.

9. La distinction entre dépression endogène et dépression réactionnelle a disparu depuis le DSM-III, après des débats assez vifs et qui mériteraient peut-être d'être repris.

10. On peut relever à ce propos que la tristesse et son éventuelle disparition n'y occupent guère de place.

11. Nous renvoyons ici aux travaux dans le domaine Science and Technology Studies, notamment Knorr-Cetina (1981), Latour (1985, 1987) ou encore Bazerman (1988). Un apport spécifique de la lecture proposée ici est de corréliser « manière de rendre compte de la preuve » et « recherche clinique ».

12. Hill est un statisticien anglais connu pour avoir dirigé ce qui est considéré classiquement comme le premier essai randomisé contrôlé sur l'efficacité de la streptomycine dans la tuberculose.

13. Le terme a été proposé par Harry Marks pour décrire les acteurs disparates d'un mouvement qui a cherché à utiliser l'expérimentation contrôlée pour « diriger la pratique médicale » selon des principes « scientifiques ».

14. « [randomization] insures that neither our personal idiosyncrasies (our likes and dislikes consciously or unwittingly applied) nor our lack of balanced judgment has entered into the construction of the different treatment groups » (Hill, 1952, p. 115).

15. « [...] we turn first to measurable characteristics [...] objective characteristics must be sought and are usually found in such features as the duration of fever, the level of the blood sedimentation rate and the presence of infecting organisms » (Hill, 1952, p. 117).

16. « [...] the judgments must be made without any possibility of bias [...] In other words, the assessment must, whenever possible, be "blind" – the assessing physician must not know the treatment of the patient » (Hill, 1952, p. 118).

17. « [...] complete ignorance in both patient and doctor of the

treatment actually given [is] essential. » (Hill, 1952, p. 118).

[18.](#) Cette formule peut légèrement varier dans la forme : trouble ou patient « remplissant les critères du DSM-III ou IV ou de l'ICD-10 », etc.

[19.](#) Par exemple : Composite International Diagnostic Interview (CIDI), Structured Clinical Interview for DSM-IV (SCID), etc.

[20.](#) Par exemple : Hamilton Depression Scale (HAM-D), Beck Depression Inventory (BDI), Young Mania Scale (YMS), Positive and Negative Syndrome Scale (PANSS), Brief Psychiatric Rating Scale (BPRS), etc.

[21.](#) On parle généralement d'échelles « validées ». Notons que les échelles psychométriques soulèvent de nombreuses questions épistémologiques que nous ne pouvons pas aborder ici.

[22.](#) Le DSM-III était un projet ambitieux, dont les auteurs espéraient qu'il aurait un impact majeur sur la clinique et la recherche psychiatrique, mais l'ampleur de son succès a surpris la profession et sans doute les auteurs eux-mêmes.

[23.](#) Si la fluoxétine n'est pas chronologiquement le premier de la classe des inhibiteurs spécifiques de la recapture de la sérotonine, elle en est le chef de file incontesté.

[24.](#) Par exemple, le Prozac® est considéré comme une « wonder drug » par le New York Magazine, qui annonce qu'on peut désormais dire « bye-bye, blues » (Schumer, 1989).

[25.](#) Relevons par exemple que 5 % des femmes américaines âgées de 45 à 64 ans ont pris un antidépresseur au cours d'un mois donné dans les années 1988-1994, contre 22 % pour la période 2005-2008 (National Center for Health Statistics, 2011).

[26.](#) Il y aura par la suite des surprises désagréables quant à l'efficacité (mise en doute en raison du biais de publication, sur lequel nous reviendrons brièvement) aussi bien que pour ce qui est des effets secondaires (avec notamment un scandale de l'olanzapine, dont le risque diabétogène a été dissimulé par le fabricant Eli Lilly, condamné à de lourdes amendes).

[27.](#) Nous traduisons.

[28.](#) La mesure d'efficacité choisie était la suivante : plus faible était le taux d'interruption ou, à défaut, plus longue était la durée jusqu'à l'interruption, meilleure serait considérée la molécule.

[29.](#) Les deux études sont différentes dans la façon dont elles sont construites, mais cherchent à répondre au même type de questions.

[30.](#) Le tableau des résultats principaux de CATIE comprend plus de 70 lignes et s'étend sur deux pleines pages, avec 11 notices explicatives quant à divers sigles utilisés (Lieberman, Stroup, McEvoy et al., 2005).

[31.](#) La méta-analyse est une méthode statistique permettant de regrouper plusieurs études (typiquement des essais randomisés contrôlés), chercher à les uniformiser et produire ainsi une sorte de super-étude dont la force réside dans le nombre (puisqu'elle s'appuie sur tous les patients inclus dans les études individuelles), au prix de difficultés méthodologiques variées. L'idée est de répondre ainsi à des

questions qui restent controversées en dépit des études déjà publiées. C'est aussi une manière commode de publier une étude sans avoir à recruter de patients, puisque l'objet de la méta-analyse est constitué d'autres études (on commence à voir paraître des méta-analyses de méta-analyses). En principe, c'est l'essai randomisé contrôlé qui est la meilleure preuve selon EBM, mais lorsque ceux-ci sont en désaccord, c'est la méta-analyse. La première méta-analyse dans Pubmed date de 1988 ; on en compte plus de 9 000 en 2012, et le nombre est exponentiel. L'étude de Cipriani emploie une méthode considérablement plus complexe qu'une méta-analyse standard.

[32.](#) Nous renvoyons le lecteur à la publication en question pour apprécier la complexité du tableau des résultats.

[33.](#) L'halopéridol est un antipsychotique classique, connu depuis les années 1960 ; la rispéridone est un antipsychotique « atypique », encore sous licence à l'époque de la parution de cette étude la comparant au lithium et à un placebo. Cette dernière étude est sponsorisée par le fabricant de la rispéridone ; néanmoins, selon les auteurs, ni l'halopéridol ni la rispéridone ne sont significativement supérieurs au lithium. Mais selon les calculs de Cipriani et al., la rispéridone et l'halopéridol sont plus efficaces que le lithium.

[34.](#) La galénique désigne l'art de composer le médicament à partir de la molécule de base : il peut exister plusieurs « formes galéniques » d'une même molécule (e.g., sirop, comprimé, suppositoire, injection intramusculaire, etc.). La galénique a un effet sur la cinétique du médicament, et on met parfois au point des « formes retard », qui permettent de réduire le nombre de prises quotidiennes.

[35.](#) « Sit down, shut up and show me the numbers » est une sorte de dicton dans les milieux économiques qui s'appliquerait sans difficulté à la recherche clinique de ces dernières années.

Nos partenaires

Le projet *Savoirs* est soutenu par plusieurs institutions qui lui apportent des financements, des expertises techniques et des compétences professionnelles dans les domaines de l'édition, du développement informatique, de la bibliothéconomie et des sciences de la documentation. Ces partenaires contribuent à la réflexion stratégique sur l'évolution du projet et à sa construction. Merci à eux !



- CONCEPTION : [ÉQUIPE SAVOIRS](#), PÔLE NUMÉRIQUE RECHERCHE ET PLATEFORME GÉOMATIQUE (EHESS).
- DÉVELOPPEMENT : DAMIEN RISTERUCCI, [IMAGILE](#), [MY SCIENCE WORK](#) DESIGN : [WAHID MENDIL](#).

P D N

Pôle Document Numérique
Maison de la Recherche en Sciences Humaines
CNRS - UNIVERSITÉ DE CAEN

métopes

méthodes et outils
pour l'édition structurée

EPFL

bnu
strasbourg

enssib

école nationale supérieure
des sciences de l'information
et des bibliothèques

CAK
Centre Alexandre-Koyré
Histoire des sciences et des techniques
UMR 8560 EHESS-CNRS-MNH
LECOLE
HAUTES
ETUDES
DE
SCIENCE
CNRS



ANHIMA

