



L'efficacité thérapeutique. La psychopathologie, de l'experimentalisme à l'évaluation

Philippe Le Moigne

► To cite this version:

Philippe Le Moigne. L'efficacité thérapeutique. La psychopathologie, de l'experimentalisme à l'évaluation. Françoise Champion. Psychothérapie et Société, Armand Colin, pp.215-235, 2008, Sociétales, 9782200350611. halshs-00573019

HAL Id: halshs-00573019

<https://shs.hal.science/halshs-00573019>

Submitted on 27 Mar 2022

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.

CHAPITRE X

L'efficacité thérapeutique *La psychopathologie, de l'expérimentalisme à l'évaluation*

Philippe Le Moigne

Centre de Recherche Psychotropes, Santé Mentale, Société (CESAMES)

La mesure de l'efficacité thérapeutique emprunte au modèle des Essais Contrôlés Randomisés (ECR). Ce modèle fait écho à une doctrine particulière de l'expérimentalisme, aujourd'hui dominante en médecine. Cette doctrine associe l'activité scientifique à la démonstration d'une action tangible sur son objet. En référence à ce principe, la recherche s'est ouverte à la mesure des effets exercés sur les troubles mentaux par les médicaments psychotropes, puis par les psychothérapies. A cette occasion, une certaine confusion s'est instaurée entre la recherche et l'évaluation des pratiques professionnelles, l'expérimentation offrant dès lors un moyen d'arbitrer les clivages théoriques et les différends techniques des psychothérapeutes. Il y a là plus qu'un paradoxe. En effet, l'évaluation des psychothérapies à partir des ECR contribue, en raison des contraintes de la méthode, à exclure la relation du thérapeute et du patient de l'étude de l'activité clinique, privant ainsi celle-ci d'une dimension essentielle à sa connaissance.

Sur le modèle de la médecine expérimentale, l'appréciation de la qualité des thérapies dédiées à la psychopathologie – des médicaments psychotropes comme des psychothérapies – s'est appuyée sur la méthode des essais contrôlés randomisés (ECR). Cette méthode, d'origine anglo-américaine, incarne une doctrine particulière de l'expérimentalisme. Elle se revendique d'une épistémologie pragmatiste de la science où prime, au-delà de la théorie ou de la conception de l'objet, la démonstration d'un effet exercé sur la matière ou le vivant. Selon cette doctrine, l'efficace fait preuve ou rend compte, à tout le moins, d'une maîtrise de son objet par la science. Parce qu'elle se réfère à ce modèle, l'évaluation des thérapies du psychisme, et des psychothérapies en particulier, n'entend donc pas caractériser le travail thérapeutique mais cherche plutôt à déterminer dans quelle mesure celui-ci inaugure bien un changement chez le patient. Deux implications se déduisent de ce projet. D'abord, l'impact thérapeutique étant placé au premier plan, les controverses relatives à l'étiologie et au traitement des psychopathologies sont supplantées, court-circuitées, sinon

réduites à la manifestation d'un pluralisme de bon aloi. Plus encore, la mesure des effets thérapeutiques peut fournir après coup une évaluation directe des mérites respectifs des cliniciens, de leurs obédiences et de leurs techniques, sans avoir à interroger leurs principes. C'est ainsi que la méthode a été promue progressivement en *un mode d'arbitrage, tant théorique, disciplinaire que professionnel*. Pour autant, et c'est là le second aspect du changement, cet expérimentalisme soumet l'étude de l'action thérapeutique à un protocole conditionné par la mise à l'écart de la relation entre le thérapeute et le patient. L'application de la méthode aux psychothérapies, thérapies fondées sur la relation s'il en est, prive donc la recherche d'une dimension essentielle à leur connaissance. A ce titre, l'application des ECR sur ce terrain est loin d'exprimer une extension logique de la méthode : elle en marquerait plutôt *le point limite*.

Comprendre ce paradoxe exige un détour par l'histoire. De fait, l'application de l'expérimentalisme à la psychopathologie est ancienne. Elle a servi des fonctions qui méritent d'être rappelées si on veut saisir la mutation que l'usage contemporain de l'expérimentation a appliquée à cette tradition. En premier lieu, l'expérimentalisme a d'abord cherché à reproduire en laboratoire les conditions d'apparition de la maladie, dans l'espoir d'en comprendre les mécanismes et de pouvoir agir sur elle. Ainsi, en pharmacologie, la recherche s'est appliquée dès le 19^{ème} siècle à reproduire les états d'hallucination à partir des propriétés psychoactives du haschisch et des opiacés en particulier, hypothèse étant faite que la physiopathologie et la thérapeutique des troubles mentaux pourraient être découvertes par ce moyen¹. Le procédé a conduit de la même manière la réflexologie, dans la lignée de Pavlov, à faire le pari d'une induction expérimentale des névroses chez l'animal. Dans le sillage du comparatisme béhavioriste, cet expérimentalisme s'est associé au projet de fonder une économie générale de la réaction d'où, pensait-on, les stimuli générateurs des comportements les plus anormaux pourraient être déduits².

En second lieu, une partie conséquente des recherches s'est appliquée à la conception et à la validation des méthodes capables de servir l'investigation de l'économie psychique. C'est ainsi que la psychologie et la psychanalyse, cherchant à accéder à la lecture interne des processus mentaux et des affections, ont entendu dès leur origine soumettre l'hypnose ou l'introspection au jeu de l'expérimentation, pour finalement en réfuter les fondements³.

Enfin, l'expérimentalisme, se fondant sur les hypothèses explicatives de la psychopathologie, s'est attaché à en valider les implications dans le champ de la nosographie et des psychothérapies. Dans cet esprit, la psychanalyse a cherché

¹ Cf. A.A. Baumeister, M.F. Hawkins, "Continuity and discontinuity in the historical development of modern psychopharmacology", *Journal of the History of the Neurosciences*, 14, 2002, pp. 199-209.

² Voir par exemple G.V. Hamilton, *An Introduction to Objective Psychopathology*, Saint-Louis, Mosby Company, 1925.

³ Cf. J. Carroy, *Hypnotisme, suggestion et psychologie. L'invention de sujets*, Paris, PUF, 1991. Sur le destin de l'hypnotisme, voir A. Mayer, « Introspective hypnotism and Freud's self-analysis procedures of self-observation in clinical practice », *Revue d'Histoire des Sciences Humaines*, 5, 2003, pp. 171-196. Pour une histoire de l'introspection en psychologie, l'une des premières méthodes expérimentées par la discipline, on pourra lire en particulier D.J. Coon, "Standardizing the subject: experimental psychologists, introspection, and the quest for a technoscientific ideal", *Technology and Culture*, 1993, 34, pp. 757-783.

très tôt, par l'intermédiaire de Carl Jung en particulier, à corroborer le caractère homogène et discret des névroses, par la recherche d'une discrimination électrophysiologique du mode émotionnel de patients⁴. Dans la même veine, des expériences psychothérapeutiques ont pu être conduites, sur la base d'hypothèses informées par la psychopathologie, dans le but de mettre à l'épreuve le mécanisme d'action prêté aux thérapies et d'en parfaire la technique. Aux Etats-Unis, au détour de la seconde guerre mondiale, ce projet animera l'ensemble des courants cliniques, depuis la psychologie humaniste de Rogers jusqu'à l'approche psychanalytique défendue par l'Ego Psychology, sans parler des thérapies comportementales où l'idée d'une telle expérimentation atteindra un point culminant⁵. La psychopathologie expérimentale s'appliquait alors à l'étude des affections et des patients, à la validation des modèles explicatifs dans la nosographie ou la cure, si bien que recherche et clinique, partageant le même objet, restaient intimement imbriquées l'une à l'autre.

A partir des années 1950, cette unité se fragilise, la clinique étant peu à peu instituée en objet de recherche. Sous ce jour, le modèle expérimental s'enrichit d'une autre vocation : il vise à déterminer l'exacte rationalité de la pratique clinique, par l'étude du jugement diagnostique des thérapeutes et la mesure du rendement atteint par le travail thérapeutique⁶. C'est ainsi que s'est fait jour aux Etats-Unis le projet d'évaluer les mérites des psychothérapies. Avec l'expérimentation, il a été dès lors question de confronter les modèles entre eux, en comparant « l'action » de la cure analytique à l'impact des techniques cognitives ou comportementales ou, plus largement encore, en mesurant les effets respectifs des psychothérapies et des chimiothérapies, en vue de départager les lectures psychogénétiques et biogénétiques de la maladie mentale⁷. Au-delà de la compréhension des troubles eux-mêmes, l'expérimentalisme se voyait ainsi chargé, plus ou moins confusément, de l'arbitrage des luttes d'écoles, du règlement des critiques adressées à l'institution psychiatrique, et de la rationalisation du diagnostic comme de la pratique thérapeutique. Cette nouvelle donne devait signer l'apparition d'un cadre inédit d'expertise, la psychopathologie expérimentale s'ouvrant à l'«évaluation clinique standardisée», selon la formule consacrée depuis.

Une telle évaluation ne pouvait que condenser une série d'enjeux, comme l'a rappelée à sa manière la récente expertise conduite en France dans ce domaine

⁴ E.I. Taylor, "Jung before Freud, not Freud before Jung: Jung's influence in American psychotherapeutic circles before 1909", *Journal of Analytical Psychology*, 1998, 43:1, 1998, pp. 97-114.

⁵ Cf. C.R. Rogers, G.Y. Gendlin, D.V. Kiesler, C. Truax (ed.), *The Therapeutic Relationship and Its Impact. A Study of Schizophrenics*, Madison, University of Wisconsin Press, 1967 ; D. Rapaport, *The Structure of Psychoanalytic Theory*, New York, International University Press, 1960 ; H. J. Eysenck (ed.), *Experiments in Behaviour Therapy*, Oxford, Pergamon Press, 1964.

⁶ Faute de place, on n'abordera pas ici le dispositif que la recherche expérimentale a conçu en vue d'évaluer la fiabilité du diagnostic en psychopathologie. Cette question est abordée dans Ph. Le Moigne, *La cause personnelle. Psychiatrie, sciences et société* (à paraître).

⁷ Le coup d'envoi de ces « tests comparatifs » est dû sans nul doute au behavioriste Hans Eysenck et à son célèbre article contre la psychanalyse où il tente de montrer que les patients névrotiques sont en droit d'attendre autant d'effets de la cure analytique que d'une rémission naturelle. H.J. Eysenck, "The effects of psychotherapy: An evaluation", *Journal of Consulting Psychology*, 16: 1952, pp. 319-324.

par l'INSERM⁸. D'abord, l'évaluation, parce qu'elle s'applique au caractère opératoire des thérapies, engage l'autorité professionnelle de la psychiatrie et de la psychologie clinique. En effet, la démonstration d'efficacité gage le crédit d'une technique, tandis que sa réfutation contribue à jeter le trouble sur sa valeur thérapeutique : logiquement, les défenseurs des modèles patentés attendent de ce juge de paix qu'il cautionne la supériorité de leur méthode, contre les modes d'intervention qui entendent à leur périphérie être admis sous le même label. C'est là, depuis la naissance de la psychologie, un fait récurrent, l'expérimentalisme ayant d'abord cherché à organiser la discipline comme science contre les modèles d'explication et de prise en charge se revendiquant d'une dimension religieuse, spiritualiste, ou bien encore du magnétisme ou de l'ésotérisme⁹. Mais, les techniques thérapeutiques étant soudées à un jeu particulièrement vif d'oppositions depuis leur origine, la comparaison de leurs mérites respectifs a également attiré contre elle les rivalités d'obédience, de disciplines et d'explications qui traversent les lectures les plus académiques de la psychopathologie, depuis la psychanalyse jusqu'aux neurosciences.

Toutefois, l'élément qui a sans doute déclenché les foudres des praticiens français a trait au peu de confiance que l'expertise des psychothérapies concède, plus ou moins implicitement, à leur compétence. Le thérapeute, dans la mesure où un droit d'exercer lui est reconnu, tend bien souvent à considérer qu'il est non seulement capable d'estimer les effets de sa pratique mais qu'il en est le seul juge. En ce sens, l'évaluation équivaut à un redoublement d'expertise : pour reprendre la terminologie des essais cliniques, elle s'applique à l'examen des « juges », auquel elle cherche à opposer un verdict plus général et plus impartial. Aux Etats-Unis, l'évaluation des thérapies s'est revendiquée de ce préalable en postulant que les psychothérapeutes n'étaient pas à même de se prononcer sur les effets de leur pratique, en raison de leurs « a priori théoriques » et de leur irréductible idiosyncrasie¹⁰. Devant cette présomption de subjectivité, les thérapeutes ont dû se soumettre à des critères d'objectivité venus d'un tout autre univers, celui des essais contrôlés randomisés. Sur quels principes repose précisément l'objectivité de ces essais ? Quel en est l'arrière-plan logique ? Comment ont-ils été appliqués à la psychopathologie et avec quel rendement ?

En vue d'apporter un début de réponse à ces questions, le texte aborde trois points en particulier. 1) Il revient d'abord sur deux des sources essentielles du dispositif d'évaluation, à savoir : le comparatisme statistique, tel qu'il a été développé en psychologie, et l'expérimentalisme clinique à proprement parler. 2) Il évoque ensuite les mesures d'adaptation qui ont dû être conçues pour permettre l'application du dispositif au terrain de la psychopharmacologie puis des psychothérapies. 3) Enfin, le texte rend compte des études que la psychologie américaine a consacrées à l'évaluation des psychothérapies.

⁸ INSERM, *Psychothérapies : trois approches évaluées*, Paris, Editions INSERM, 2004. Comme on le sait, cette expertise des psychothérapies a connu un destin pour le moins controversé.

⁹ Cf. D.J. Coon, "Testing the limits of sense and science: American experimental psychologists combat spiritualism, 1880-1920", *American Psychologist*, 47 (2), 1992, pp. 143-151 ; T. Dehue, "From deception trials to control reagents: The introduction of the control group about a century ago", *American Psychologist*, 55, 2000, pp. 264-269.

¹⁰ Sous l'influence en particulier de Robert Spitzer, le rédacteur en chef du DSM-III. R. Spitzer, J. Cohen, J. Fleiss, J. Endicott, "Quantification of agreement in psychiatry diagnosis: A new approach". *Archives of General Psychiatry*, 17, 1967, pp. 83-87.

1. Les principes de l'évaluation : aux origines des essais cliniques

Les essais cliniques, qu'ils s'appliquent aux pathologies somatiques ou psychiques, procèdent de la manière suivante : ils établissent l'effet attribuable à la thérapie en comparant le degré d'amélioration de patients atteints du même trouble, selon qu'ils ont bénéficié du traitement ou reçu un placebo, voire encore un thérapie de substitution. L'allocation des patients dans chacun des groupes, « traités » et « témoins », se fait de manière aléatoire (par randomisation), et à l'insu tant des sujets que des cliniciens chargés d'évaluer leur rémission. Les fondements auxquels puise cette procédure sont hérités de la statistique, psychologique en particulier, et d'une forme particulière de l'expérimentalisme médical. On les présente à grands traits.

Contrôle et randomisation : le comparatisme statistique en psychologie

L'une des premières sources des essais cliniques est liée au développement de l'expérimentation en sciences humaines et, à son corollaire, la création de la statistique. Au milieu du 19^{ème} siècle, l'Etat a cherché à intervenir sur les lois du marché en se prévalant de mesures efficaces, et suffisamment impartiales. Avec le développement des sciences humaines, la décision publique entendait ainsi s'appuyer sur un jugement plus neutre et plus impersonnel que celui de l'érudit. Au début du 20^{ème} siècle, les premières échelles standardisées, dédiées à la mesure de la délinquance ou de la pauvreté, font leur apparition. En effectuant une mesure sur un territoire donné, avant et après l'engagement d'un programme d'action, on pensait ainsi être capable d'évaluer l'impact de la décision publique sur la vie sociale. Toutefois, très vite, la question s'est posée de savoir quelle part du changement observé pouvait être imputée en propre à l'intervention planifiée. Aux Etats-Unis, le sociologue F.S. Chapin y a vu une impossibilité : l'idée d'évaluer l'effet d'une mesure publique de façon purement « étanche » était, selon lui intenable, en vertu de la multiplicité des facteurs agissant sur la vie sociale¹¹. L'origine ethnique, le statut socio-économique, les opinions, etc., constituaient autant d'éléments susceptibles d'interférer avec le phénomène mesuré. L'expérience ne pouvait pas être contrôlée. En effet, on ne pouvait pas constituer deux groupes parfaitement similaires, dont l'un aurait été soumis à la mesure préconisée par l'action publique tandis que l'autre aurait servi de « contrôle ». Et, selon Chapin, même si la sociologie en avait été capable, l'opération aurait abouti à créer des groupes parfaitement abstraits, de sorte que la mesure de l'action publique aurait été renvoyée à une sorte de fiction sociale¹².

¹¹ F.S. Chapin, *Experimental Design in Social Research*, New York, Haper. 1947.

¹² On trouvera chez Durkheim le même argument : pour lui, l'expérimentation ne pouvant être qu'indirecte en sociologie, celle-ci devait d'abord recourir à l'analyse comparée des données laissées comme traces par les faits sociaux, plutôt que de s'attacher à les construire. E. Durkheim, *Les règles de la méthode sociologique*, Paris, Alcan, 1895. Pour un point de vue analogue, voir M. Halbwachs, « L'expérimentation statistique et les probabilités », *Revue Philosophique*, 96, 1923, pp. 340-371.

C'est la psychologie de l'éducation qui a introduit l'idée de comparer des groupes artificiels plutôt que naturels. L'évaluation comparée des méthodes pédagogiques butait sur la même difficulté : comment constituer des groupes d'études homogènes, sachant que l'impact des techniques éducatives a partie liée avec le niveau intellectuel, l'origine sociale ou bien encore le genre de l'élève ? Dans les années 1920, W. McCall recommande d'utiliser le hasard en vue de composer les deux populations de recherche, sous réserve que l'échantillon soit assez important : l'affectation d'un élève par tirage au sort, soit dans le « groupe testé » (soumis à la technique évaluée), soit dans « le groupe témoin », permettait de garantir en probabilité la stricte équivalence des populations composées par ce moyen. Avec McCall, l'expérimentalisme devenait inséparable de la randomisation, c'est-à-dire de l'affectation aléatoire des sujets dans les groupes d'évaluation¹³.

Dans les années 1930, le statisticien Ronald Fisher conclut, par d'autres voies, à la même nécessité de composer la population d'étude par allocation aléatoire. Surtout, en vue de déterminer, à l'issue de l'épreuve, dans quelle mesure les groupes font part de différences significatives, il élabore un test statistique resté fameux : l'analyse de co-variance¹⁴. Mais, Fisher prend soin d'indiquer quelle est la valeur statistique de ce genre d'épreuve. L'analyse de co-variance permet d'établir qu'un changement significatif a bien eu lieu entre t1 et t2 dans la population traitée, comparativement à celle qui ne l'a pas été, mais elle est incapable d'en établir la cause. Rapporté au domaine médical, un tel principe revient à admettre que la statistique ne permet pas de produire une étiologie ou une interprétation des effets thérapeutiques. Elle exprime une mesure de différenciation, de changement, mais elle n'a pas le pouvoir d'en déterminer l'origine¹⁵. On tient avec elle un constat non pas une preuve. Pour parler de démonstration, il faut être en mesure de relier l'évolution attestée par la statistique à l'énoncé d'un principe actif ou d'une pathophysiologie, en un mot, à une explication. A cet égard, l'introduction des essais cliniques marque bien l'avènement en médecine d'une organisation de recherche de plus en plus diffractée. Elle accompagne une spécialisation croissante des domaines d'investigation, orchestrée par une division de plus en plus ferme entre l'étiologie, la thérapeutique, la sémiologie des maladies et la clinicométrie elle-même.

Il est vrai que, chez Fisher, l'expérimentation s'était d'abord appliquée à l'agronomie, c'est-à-dire au rendement acquis par l'adjonction d'engrais dans la production agricole. Dans ce cadre, la randomisation n'affecte pas la population randomisée. En revanche, l'application du procédé à l'être humain doit composer avec l'interaction sociale et, en médecine en particulier, avec la relation du praticien et du patient. C'est là la raison d'être du placebo et de la méthode dite du « double aveugle » qui viendront compléter les dispositifs des essais contrôlés randomisés, l'une et l'autre de ces dispositions cherchant à répondre à la question suivante : comment isoler l'action des thérapeutiques de la relation thérapeutique elle-même ?

¹³ W.A. McCall, *How to Experiment in Education*, New York, Macmillan, 1923.

¹⁴ R. Fisher, *The Design of Experiment*, Edinburgh, Oliver and Boyd, 1935.

¹⁵ Cf. R. Ashcroft, "What is clinical effectiveness?", *Studies in History and Philosophy of Biology and Biomedical Sciences*, 33, 2002, pp. 219-233.

Une certaine approche de l'expérimentation médicale : aux origines du placebo et du « double aveugle »

L'introduction du placebo est très ancienne, même si elle n'a été associée à la randomisation qu'au 20^{ème} siècle. Le procédé, introduit au 18^{ème} siècle, s'appuyait sur le constat suivant : la pratique médicale avait révélé que des substances inactives pouvaient être porteuses d'amélioration, en vertu d'effets dits « psychologiques », assez mal identifiés¹⁶. En distribuant le traitement et le placebo à l'insu des patients (c'est-à-dire alors en « simple aveugle », le praticien ayant connaissance de la distribution du traitement), la procédure laissait présager que l'amélioration procurée par le traitement, si elle s'avérait significative, pourrait être affectée à la part proprement active de la thérapeutique, et non à un pur effet de suggestion¹⁷. Ce fût alors, entre autres, une manière d'organiser la défense de la médecine allopathique contre l'homéopathie¹⁸.

La perspective a soulevé de nombreuses questions éthiques : avec elle, le thérapeute a dû admettre de ne pas traiter certains des patients, tout en leur laissant supposer qu'éventuellement il les traitait bel et bien¹⁹. A partir des années 1950, l'expérimentalisme médical, qu'on désigne aujourd'hui plus communément sous le terme de médecine des preuves – *Evidence-Based Medicine*, va ajouter l'idée d'étendre cette clause au médecin expérimentateur afin qu'il ne sache pas si le sujet qu'il examine est ou non traité. Avec cette extension, la procédure expérimentale allait prendre la forme qu'on lui connaît aujourd'hui, à savoir, celle de l'essai clinique randomisé avec placebo en double aveugle²⁰. Un tel dispositif, construit à l'insu du clinicien, cherche à l'évidence moins à tirer parti de ses différents savoirs qu'à les mettre à l'écart. Pourquoi chercher ainsi à maintenir le thérapeute dans l'ignorance du protocole de recherche ?

Cette ambition est là encore fort ancienne. Elle précède la médecine expérimentale de Claude Bernard et s'en distingue par son projet. Chez Bernard, qui par ailleurs exècre la statistique, l'expérimentation est toujours précédée d'une hypothèse déduite d'une observation. Même si chez lui l'expérience a bien pour objet d'agir sur le vivant, et ses développements morbides, elle reste néanmoins

¹⁶ L'effet placebo reste d'ailleurs aujourd'hui encore très mal connu. On estime par exemple qu'il serait plus grand lorsqu'il est question de tester une substance morphinique que de l'aspirine, l'explication de ce phénomène étant attribué au fait que les patients anticiperaient davantage d'effets dans le premier cas que dans le second. Pareille conjecture, si elle s'avérait exacte, obligerait à évaluer, pour chaque médicament, la part exacte de l'effet placebo en vue d'établir la part proprement active du produit testé. Le débat a fait rage dans l'examen des propriétés actives des antidépresseurs de dernière génération, le gain thérapeutique des substances étant très faiblement supérieur à celui obtenu du placebo. Cf. I. Kirsch, "Specifying non-specifics: Psychological mechanisms of placebo effects", in A. Harrington (ed.), *The Placebo Effect: An Interdisciplinary Exploration*. Cambridge, Harvard University Press, 1997, pp. 166-186.

¹⁷ Rappelons-le, le développement des psychothérapies s'est fondé initialement sur l'effet de suggestion que le thérapeute était censé exercer sur le patient, au moyen de l'hypnose en particulier. Pour un exemple parmi d'autres, tiré de la littérature américaine, voir B. Sidis, *Psychology of Suggestion*, New York, Appleton, 1898.

¹⁸ Cf. T.J. Kaptchuk, "Intentional ignorance. A history of blind assessment and placebo controls in medicine", *Bulletin of the History of Medicine*, 72, 1998, pp. 389-433 ;

¹⁹ Cf. H. Marks, *La médecine des preuves. Histoire et anthropologie des essais cliniques*, Le Plessis-Robinson, Institut Synthélabo, 1999.

²⁰ Pour une présentation canonique de la procédure, voir S.J. Pocock, *Clinical Trial. A Practical Approach*, New York, John Wiley, 1983.

guidée par un principe explicatif ou, à tout le moins, par une conjecture²¹. L'expérimentalisme médical contemporain, tel qu'il est soudé à la méthode des essais cliniques, doit son origine à une perspective plus radicale qui exclut l'apport de la théorie comme celui de l'intelligence pratique, pour ne considérer que l'impact que la recherche est capable d'exercer sur le vivant²².

Si on se réfère à Ulrich Tröhler, ce projet naît également au 18^{ème} siècle²³. La médecine est alors divisée entre deux approches : le rationalisme et l'empirisme. La première place le savoir médical dans la théorie, c'est-à-dire dans une connaissance raisonnée de la pathologie et des causes. La seconde privilégie la pratique médicale et l'art thérapeutique. Avec elle, prédominent l'expérience clinique, le contact avec le patient, l'habitude de prescription, l'examen et le traitement individualisés des malades, etc. Le problème rencontré par le rationalisme médical a trait à son degré d'abstraction : il n'est pas en mesure d'estimer l'efficacité des thérapeutiques, pas plus qu'il ne peut juger de la variation de leurs effets d'un individu à l'autre. Quant à l'empirisme, s'il est bien guidé par l'intervention médicale, il est incapable d'indiquer pourquoi, au-delà du constat d'expérience, une pratique clinique mérite d'être généralisée. Les deux versants, théoriques et pratiques, de la médecine faisaient ainsi face à la même part d'ignorance. L'expérimentation, alors naissante, a cherché à faire de cette difficulté un atout. Selon elle, si l'ignorance était utilisée de façon intentionnelle, et gérée de manière méthodique et réglée, elle pouvait être à la source d'un savoir nouveau²⁴. Pour cela, il importait que l'expérience scientifique fasse abstraction des savoirs théoriques et de l'art médical, et qu'elle réduise à son plus bas niveau l'individualité du patient et la variabilité formelle de la pathologie. Privé ainsi de ses projections de sens, le médecin chargé d'examiner l'évolution du malade n'aurait plus eu, selon le modèle, qu'à considérer les faits.

La stratégie du double aveugle, qui s'impose à partir de 1950, n'a fait que systématiser ce principe. En effet, cette stratégie entend, de la même manière, faire du praticien un pur observateur, dépouillé de toute raison théorique et professionnelle. En optimisant son ignorance, la procédure vise à se cantonner au donné, et à s'assurer de la neutralité et de l'indépendance de ses observations. Le plan de répartition des groupes étant seul connu du statisticien, l'essai clinique

²¹ C. Bernard, *Principes de médecine expérimentale. Ou de l'expérimentation appliquée à la physiologie, à la pathologie et à la thérapeutique*, Paris, PUF, 1947 [Ecrits de 1858 à 1877]. Pour une lecture desdits principes, se reporter à M.D. Grmek, *Le legs de Claude Bernard*, Paris, Fayard, 1997.

²² On reconnaîtra ici l'empreinte de la philosophie des sciences de Bacon. Celle-ci n'a pas seulement essaimé en médecine. Elle irrigue les accents les plus prométhéens de la psychologie béhavioriste, de Skinner en particulier (Cf. L.D. Smith, « On prediction and control: B.F. Skinner and the technological ideal of science », *American Psychologist*, 40 (2), 1992, pp. 216-223.). Par certains aspects, la première épistémologie défendue par Ian Hacking procède de l'extension de cette philosophie à la physique. Avant d'opter pour une posture plus constructiviste, il considérait en effet que l'activité scientifique s'éprouve d'abord à l'intervention et à la manipulation que la recherche est capable d'opérer sur la matière, les représentations de l'objet et de ses processus n'intervenant qu'à la marge. I. Hacking, *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science*, Cambridge, Cambridge University Press, 1983.

²³ U. Tröhler, *To Improve the Evidence of Medicine. The 18th Century British Origins of a Critical Approach*, Edinburgh, Royal College of Physicians, 2000.

²⁴ *Ibid.* Voir également T.J. Kaptchuk, "Intentional ignorance. A history of blind assessment and placebo controls in medicine, *art. cit.*

paraissait ainsi s'être prémuni des biais susceptibles d'affecter son déroulement : l'idiosyncrasie du médecin, les obédiences d'école, les éventuelles collusions avec les laboratoires pharmaceutiques étant neutralisées, seule la symptomatologie et son évolution seraient prises en compte.

2. L'application du modèle à la psychopathologie

La psychopathologie quantitative, nom aujourd'hui donné à la psychiatrie fondée sur la statistique, a repris à son compte les termes du dispositif imaginé par la statistique et l'expérimentalisme médical ; elle en a fait sa raison, il est vrai, non sans avoir procédé à certains ajustements²⁵. Mais, avant d'évoquer la teneur de ces aménagements, une question qui reste souvent plus inaperçue mérite d'être abordée : comment les auteurs des essais cliniques ont-ils conçu d'apprécier l'amélioration des patients atteints d'un trouble psychisme, étant considéré qu'un examen biologique ou physiologique ne peut ici, à la différence du cas présenté par les pathologies somatiques, valoir comme critère d'évaluation ?

La caractérisation de l'évolution

Le cadre logique de l'essai clinique marque un resserrement de l'analyse sur l'évolution de la pathologie. Dans la mesure où prime ici la démonstration du changement d'état des patients, l'étiologie ou les propriétés des thérapies ne sont considérées qu'à un titre subalterne et laissées à l'examen d'autres secteurs de recherche. Seule importe la métrologie du changement. Ce plan de travail a été instauré dès l'invention des médicaments psychotropes. Aux Etats-Unis, le Psychopharmacological Service Center, qui préfigure la naissance de la psychopharmacologie, se rallie à ce découpage de la preuve dès 1957²⁶. Outre-atlantique, l'amendement de 1962, qui fait suite à la mise en évidence des effets tératogènes de la thalidomide, donnera à cette division du travail le caractère d'une exigence légale, les laboratoires pharmaceutiques devant dès lors assurer l'Etat de l'innocuité comme de l'efficacité des thérapeutiques qu'ils proposent à la commercialisation. Toutefois, les modalités de cette évaluation ne seront stabilisées pour la psychopathologie quantitative qu'au milieu des années 1970, sous l'influence en particulier du manuel produit par la Early Clinical Drug Evaluation Unit du National Institute for Mental Health²⁷. Cette durée de gestation pourrait être imputée aux difficultés qui laisse augurer en propre l'examen de l'évolution des troubles mentaux. Ce serait faire abstraction de la bataille que le règlement de cette question a suscitée entre psychiatres et psychologues. Cette bataille, engagée sur le terrain de la composition des échelles cliniques, a condensé une série de rivalités à la fois 1) professionnelles, 2) conceptuelles, 3) et instrumentales.

²⁵ A ce sujet, voir J.D. Guelfi., V. Gaillac., R. Dardennes R. (dir.), *Psychopathologie quantitative*. Paris, Masson, 1995.

²⁶ Cf. D. Healy, *The Antidepressant Era*, Harvard, Harvard University Press, 1997.

²⁷ W. Guy, *Early Clinical Drug Evaluation - ECDEU Assessment Manual for Psychopharmacology*, Rockville, National Institute of Mental Health, 1976.

1) Au début des années 1950, R.D. Wittenborn, produit l'une des premières grilles d'évaluation dédiées à la psychopathologie²⁸. Ce psychologue psychométricien cherche à livrer un outil diagnostique. Mais sa grille d'observation s'applique également à l'évolution de la pathologie : l'examen répété du patient, par la passation à chaque fois réitérée de l'instrument, permet d'obtenir une cotation de la maladie dans le temps, et donc une mesure de son éventuelle rémission. L'outil est rejeté par les psychiatres opérant dans le champ de l'instrumentation clinique. Ils considèrent que le même instrument ne peut œuvrer à la fois au diagnostic et à l'évaluation du changement. Leur argument est simple : lorsqu'on cherche à évaluer l'une et l'autre de ces dimensions à la fois, on ne sait plus distinguer les entités cliniques de leurs états rémittents. Autrement dit, indépendamment de la mesure d'évolution, il importait que les critères de la pathologie soient définis et fixés à l'avance au même seuil pour l'ensemble des patients intégrés à l'épreuve. Le diagnostic et l'appréciation du changement devaient donc, selon ce principe, faire l'objet d'une division du travail, elle-même assortie d'une spécialisation instrumentale. Evidemment, en défendant ainsi la nécessité de distinguer l'instrumentation nosographique des tests appliqués à l'évaluation thérapeutique, les psychiatres cherchaient tout également à préserver, contre les psychologues, leur prérogative sur la caractérisation des pathologies mentales.

2) Le second point de friction est né, précisément, d'un désaccord sur la manière de conceptualiser le trouble mental. En 1962, John Overall publie sa Brief Psychiatric Rating Scale²⁹. Il est directeur en chef du Bureau des Standards de la très influente Veteran Administration qui s'emploie au soin des militaires revenus du front et qui, à ce titre, a fourni à la psychologie clinique l'un de ses premiers réservoirs d'emplois. En bon psychométricien, Overall est dimensionnaliste. En effet Spearman, l'initiateur de la spécialité, considérait, au-delà des tests d'intelligence qu'il avait pu créer, que les facultés psychologiques étaient étagées à la manière d'une distribution normale³⁰. En vertu de cet argument, l'aptitude, la motivation puis après la personnalité ont été considérées comme des universaux, partagés par l'ensemble de la population, ne variant qu'en degrés d'un individu à l'autre. En référence à ce modèle, Overall a considéré qu'il était possible d'étager la pathologie mentale, depuis ses occurrences les plus frustes jusqu'à ses formes les plus sévères, sur un simple continuum ; dans cette hypothèse, les névroses seraient placées sur le gradient le plus proche de la normalité, tandis que les psychoses seraient positionnées au sommet de l'échelle par leur gravité. Ainsi bâti, son instrument pouvait en probabilité servir au diagnostic de n'importe quelle pathologie. La dimension générique de l'outil a suscité l'opposition des psychiatres tenants d'une psychopathologie quantitative plus respectueuse de la tradition nosographique. Cette tradition, dominée par la classification des maladies de Kraepelin, est catégorialiste. Les entités cliniques sont comprises comme des classes distinctes. Elles relatent le fait que, chez

²⁸ Cf. J.R. Wittenborn, *Wittenborn Psychiatric Rating Scales Manual*, New York, Psychological Corporation, 1955.

²⁹ J.E. Overall, D.R. Gorham, "The Brief Psychiatric Rating Scale", *Psychological Report*, 10, 1962, pp. 799-812.

³⁰ C. Spearman, *The Abilities of Man. Their Nature and Their Measurement*, New York, Macmillan, 1927.

certaines groupes d'individus, une série particulière de symptômes est réunie avec une fréquence qui ne peut être imputée au hasard. Dans la mesure où c'est ici la spécificité des agrégats de symptômes qui permet de discriminer la maladie et de différencier les entités entre elles, la psychopathologie ne peut en aucune façon être décrite sous la forme d'un continuum. En vertu de ce différend conceptuel, l'échelle d'Overall continuera d'être utilisée dans l'évaluation des neuroleptiques, où elle tient lieu de standard, mais ne pourra jamais prétendre au diagnostic des schizophrènes.

3) Il fut ainsi établi peu à peu que l'instrumentation diagnostique devait s'appuyer sur la nosographie psychiatrique. En revanche, le projet de mesurer l'évolution du patient n'avait pas encore trouvé de réponse. Les tests imaginés par les psychologues pouvaient se destiner à cette spécialisation mais, pour les psychiatres, ils comportaient un écueil : ces questionnaires étaient auto-administrés, c'est-à-dire renseignés par le patient³¹. La généralisation de la procédure à l'évaluation de la schizophrénie était difficile. Surtout l'auto-évaluation sapait l'autorité des psychiatres, en déléguant l'examen de la morbidité au malade lui-même. Par contraste, la psychiatrie entendait faire valoir sur ce terrain tout l'intérêt de l'interrogatoire médical. Cette opposition de perspectives a pu bénéficier d'un règlement instrumental : l'évolution des patients allait être mesurée à partir des tests des psychologues, mais ceux-ci seraient administrés par « hétéro-évaluation », sur la base des observations effectuées par le clinicien.

Ce découpage instrumental, entre échelle diagnostique et échelle d'appréciation du changement, marque un *modus vivendi*. D'un côté, les psychiatres ne pouvaient tout à fait se passer des psychologues dans la mesure où l'évaluation standardisée emprunte une part essentielle de son langage et de sa méthode à la psychométrie. De l'autre, la division du travail entre les évaluations diagnostiques et thérapeutiques permettait à la psychométrie de résoudre une aporie de fond : la conceptualisation du changement lui-même. En effet, rien dans le corpus théorique ou instrumental de la psychométrie ne permet d'aborder cette question. Comme l'a souligné Robert Cattell, l'une des plus fameux psychométriciens d'après-guerre, les objets étudiés par la spécialité doivent être considérés comme autant d'invariants personnels, de prédispositions structurelles, en un mot, comme autant de *traits*. Dans ce cadre, le changement ne peut être pensé que sous la forme de dispositions temporaires, indiquant qu'une transformation significative de la personnalité est en cours. Ces épisodes de transformation prennent chez Cattell le nom d'*états*, et se manifestent en particulier par une modification importante de l'humeur³². Or, selon la psychopathologie, le trouble de l'humeur est le signe même d'un épisode dépressif. Autrement dit, le raisonnement psychométrique, si on devait le traduire dans les termes de la symptomatologie psychiatrique, reviendrait à assimiler la

³¹ C'est, assez paradoxalement, l'une des raisons qui a présidé jusqu'au milieu des années 1970 au succès de l'échelle de Zung dans l'évaluation de la dépression. L'instrument comprenant un inventaire destiné à l'examen diagnostique, et un test rempli par le patient, permettait de croiser le jugement de l'expert à celui du malade. En psychiatrie, une fois que l'évaluation standardisée du diagnostic a été formellement définie et consacrée comme objective, cette perspective n'a plus compté. W.W.K. Zung, "The depression Status inventory: an adjunct to the self-rating depression scale", *Journal of Clinical Psychology*, 28, 1972, pp. 539-543.

³² R.B. Cattell, P. Kline, *The Scientific Analysis of Personality and Motivation*, London, Academic Press, 1977.

notion d'état psychologique au syndrome dépressif, et à faire de celui-ci le signe avant-coureur d'une recomposition personnelle. Evidemment, cette réunion de sens ne recouvre d'aucune manière la caractérisation que la psychiatrie a cherché à donner de la dépression ou bien encore des troubles de la personnalité.

Pour tenter d'apporter un règlement à la question du changement, l'évaluation a cherché à lier plus ou moins explicitement les approches catégorielles et dimensionnalistes. Dans ce cadre, les psychiatres ont été appelés à produire les échelles diagnostiques qui devaient permettre de sélectionner les patients des essais cliniques : ces échelles présentent, pour chacun des troubles recensés par la nosographie, la liste des symptômes requis pour qu'un diagnostic positif puisse être retenu³³. De leur côté, les tests des psychologues, reconstruits sous la forme d'échelles se prêtant à l'hétéro-évaluation, ont été réorganisés en vue de satisfaire bon an mal an aux exigences d'une mesure du changement. A cette occasion, l'approche psychométrique, acquise à la gradualité de la pathologie mentale, a cédé la place à une pure mesure de l'intensité morbide. Dans le cadre de cette métrologie, chaque échelle doit se destiner à une entité clinique et une seule. Celle-ci est réduite à une liste de symptômes auxquels est affecté un coefficient. La mesure de l'intensité du trouble procède du cumul des cotations que le clinicien établit pour chaque symptôme, après examen du patient³⁴. Ce chiffrage est entamé lors de la mise en oeuvre de l'essai clinique, puis est répété périodiquement jusqu'à son terme, en vue d'obtenir un état quantifié de l'évolution de la pathologie sous thérapeutique, et vaut ainsi comme indicateur du changement clinique³⁵.

Le protocole d'évaluation des psychothérapies

La division du travail d'évaluation entre nosographes et psychométriciens a fortement marquée l'organisation des essais cliniques dédiés aux chimiothérapies. L'essentiel de son esprit a été conservé lorsque l'évaluation s'est appliquée, au-delà des études de psychopharmacologie, à la mesure d'efficacité des psychothérapies. Deux raisons au moins ont présidé à l'extension de la recherche aux psychothérapies. Il s'est d'abord agi de proposer les thérapies au remboursement des soins, sur la base d'une efficacité démontrée³⁶. Il s'est agi

³³ C'est là la vocation des échelles dites « critères » auxquelles emprunte directement le DSM-III. Les auteurs du Manuel se sont inspirés de deux instruments qu'ils avaient composés avant lui : les « Critères de Feighner » et les « Critères de Recherche Diagnostique ». Cf. J.P. Feighner, E. Robins, S.B. Guze, R.A. Woodruff, G. Winokur, R. Munoz, "Diagnostic criteria for use in psychiatric research", *Archives of General Psychiatry*, 1972, 26, pp. 57-63 ; R.L. Spitzer, J. Endicott, E. Robins, "Research Diagnostic Criteria: Rationale and reliability", *Archives of General Psychiatry*, 1978, 34, pp. 773-782.

³⁴ Ces échelles ont données lieu elles-mêmes à des standards, tels que l'échelle d'Hamilton pour l'évaluation dynamique des états dépressifs. Cf. M. Hamilton, "A rating scale for depression", *Journal of Neurology, Neurosurgery and Psychiatry*, 23, 1960, pp. 56-62.

³⁵ C'est au Français Pierre Pichot qu'on doit la présentation la plus achevée, et sans doute la plus orthodoxe, de cette répartition instrumentale des tâches. Cf. P. Pichot, R. Olivier-Martin (eds.), *Psychological Measurements in Psychopharmacology*, Basel, S. Kruger, 1974. Voir également, J.D. Guelfi., *L'évaluation clinique standardisée en psychiatrie*, Rueil-Malmaison, Editions Pierre Fabre, 1998.

³⁶ Sur ce point, voir W.E. Barton, *The History and Influence of the American Psychiatric Association*, Washington, American Psychiatric Press, 1987.

ensuite, non plus de comprendre le fonctionnement des psychothérapies à proprement parler, mais d'en certifier l'efficacité par un examen statistique³⁷. Pourquoi ? Les thèses de la psychopathologie, nourries par l'étude de cas et le témoignage du thérapeute, avaient perdu leur autorité devant les coups de boutoir assénés entre autres par l'antipsychiatrie et sa dénonciation de l'arbitraire diagnostique en psychiatrie. Dans ce cadre, la clinique ne pouvait prétendre retrouver une légitimité qu'en recourant au mode d'objectivité et d'interprétation le plus dépersonnalisé et délocalisé qui soit, à savoir, celui des nombres³⁸.

Ici, toutefois, la rigueur de l'experimentalisme a dû concéder certains aménagements de procédure. D'abord, dans ce cas, l'acte thérapeutique ne peut être mené à l'insu du thérapeute, ni être caché au patient : la psychothérapie réclame la participation active des deux acteurs du traitement et ne peut par conséquent être soumise à la règle du double aveugle. Cet amendement a placé l'évaluation devant deux problèmes inédits. En premier lieu, les cliniciens et les patients étant engagés dans le processus thérapeutique en toute connaissance de cause, à qui reviendrait le droit d'apprécier l'efficacité du traitement ? La procédure de notation revenait à requérir du praticien qu'il évalue son propre travail, et du patient qu'il se prononce sur la compétence de son thérapeute. Le recours à un juge indépendant ne laissait pas augurer un meilleur arbitrage puisque, dans pareil cas, l'évaluation des mérites du clinicien serait laissée à un confrère.

En second lieu, à la différence des thérapeutiques médicinales qui peuvent circuler sans que leur structure moléculaire en soit changée, le format des psychothérapies est nécessairement appelé à varier : il évolue d'un thérapeute à l'autre, et reste marqué par l'échange avec le patient. Comment mesurer l'effet d'un événement si celui-ci n'est pas stable ? Pour faire face à ce problème, les évaluateurs ont cherché à soumettre à l'épreuve des faits des techniques strictement homogènes en les calibrant au moyen de protocoles. La standardisation des techniques évaluées paraissait d'autant plus requise qu'en l'absence du double aveugle on pouvait craindre une variabilité massive de l'exercice thérapeutique, et le retour en force des biais imputés aux convictions du thérapeute ou à l'individualité du patient. Mais, en soumettant le déroulement des thérapies à un tel paramétrage, l'expérimentation s'est appliquée à l'examen d'une clinique dont on peut douter qu'elle corresponde à l'exercice réel de la psychothérapie.

Enfin, dans la mesure où l'existence de « psychothérapies inertes » est difficile à concevoir, l'usage du placebo, indissociable de la définition canonique des essais randomisés, a dû laisser place à deux autres formules : soit le groupe témoin bénéficie d'une psychothérapie concurrente à celle qui est testée en

³⁷ C'est à Paul Meehl, psychanalyste et psychométricien, qu'on doit l'une des premières défenses argumentées de ce principe. P.E. Meehl, *Clinical versus Statistical Prediction. A Theoretical Analysis and a Review of the Evidence*, Minneapolis, University of Minnesota Press, 1954.

³⁸ Sur le caractère impersonnel prêté à la statistique, et sur l'influence qu'une telle attribution de sens a exercée sur le développement de la spécialité dans les sciences, et la psychologie en particulier, voir G. Gigerenzer, "Probabilistic thinking and the fight against subjectivity", in L. Krüger, G. Gigerenzer, M. Morgan, (eds.), *The Probabilistic Revolution. Ideas in the Sciences*, Vol. 2, Cambridge, MIT Press, 1987, pp. 49-72 ; T.M. Porter, 1992, "Objectivity as standardisation. The rhetoric of impersonality in measurement, statistics and cost-benefit analysis", *Annals of Scholarship*, 9, 1992, pp. 19-59.

première intention, soit il est composé de patients qu'on inscrit sur une liste d'attente, en leur promettant qu'un traitement leur sera délivré à échéance de 4 ou 6 mois. A vrai dire, par ce moyen, la procédure cherche plutôt à apprécier la proportion des rémissions naturelles qu'à isoler la part de suggestion imputable à la relation thérapeutique. Et on comprend pourquoi : il serait en effet assez curieux de vouloir isoler la part redevable à « l'effet relationnel » dans le cas de thérapies qui font précisément de la relation le principe de leur action. Autant dire que la méthodologie des essais cliniques semble buter avec l'évaluation des psychothérapies sur un point limite. Mais, voyons en quoi plus précisément.

3. La méthodologie des essais devant la relation clinique

Un artifice de procédure

Que disent les évaluateurs eux-mêmes sur la façon dont les essais cliniques mettent à l'épreuve les psychothérapies ? On peut se référer ici à l'article de synthèse que Sol Garfield a consacré à cette question dans un livre publié par l'American Psychological Association³⁹. A travers cinq points, Garfield montre combien il est difficile d'assimiler les essais randomisés à une mesure même approchée de la pratique clinique. Voici son argument :

1) Afin de s'assurer que les expérimentateurs se réfèrent bien à la même technique thérapeutique, ceux-ci doivent suivre à la lettre le protocole d'intervention consigné dans le manuel d'expérimentation. Cette « manualisation » de la clinique, selon l'expression de Garfield, produit les termes d'un exercice thérapeutique qui n'existe pas. Il s'agit d'une pratique experte de la psychanalyse, des thérapies cognitives ou comportementales qui entend incarner un modèle d'activité thérapeutique mais dont on n'est incapable de dire s'il informe de quelque façon que ce soit l'exercice quotidien du thérapeute. C'est pourquoi, selon Garfield, les conclusions émises à propos de l'efficacité d'une technique, à partir des essais cliniques, ne peuvent pas être généralisées au-delà de ce cercle.

2) L'évaluation met l'accent sur la technique thérapeutique, mais fait complètement abstraction du thérapeute. Or, les thérapies expérimentées rencontrent surtout l'adhésion des psychologues universitaires et des chercheurs eux-mêmes : 40% d'entre eux revendiquent une souscription directe aux modèles cognitifs ou comportementaux, contre 10% environ des praticiens. Ceux-ci avouent plutôt, dans leur très grande majorité, se ranger à une posture éclectique, combinant plusieurs modèles à la fois. Autrement dit, il y a loin entre les techniques évaluées et la pratique clinique de la majorité des thérapeutes.

3) Les patients sélectionnés au moyen des échelles diagnostiques sont, en raison des règles d'inclusion des essais, peu représentatifs des cas examinés par la clinique. Ainsi, en vertu de ces critères (sévérité des symptômes, absence de comorbidité, de prescription en cours, de risque suicidaire, etc.) on a pu estimer, par exemple, que seuls 10% environ des déprimés pourraient en probabilité intégrer

³⁹ S.L. Garfield, "Major issues in psychotherapy research", in D. Freedheim (ed), *History of Psychotherapy. A Century of Change*, Washington, APA, 1992, pp.335-359.

les essais dédiés à la psychopharmacologie de la dépression⁴⁰. En matière de psychothérapie, le biais de sélection est d'autant plus grand que les sujets sont volontaires, et donc déjà largement acquis au bénéfice potentiel de ce type de thérapies. C'est pourquoi, selon Garfield, il y aurait lieu de se demander ce qu'il adviendrait des mesures d'efficacité si on entendait les essais aux publics les moins enthousiastes à l'égard des psychothérapies.

4) La distinction expérimentale des groupes témoins et des groupes contrôle ne peut être assurée complètement dans les faits. Pour les patients en liste d'attente, la promesse d'un traitement différé ne suffit pas toujours à calmer leur plainte, parfois, elle y ajouterait plutôt. Ici « ne rien faire » relève à bien des égards d'une action qui influe sur l'évolution de l'affection présentée par le patient. De la même manière, la répartition au hasard des sujets est souvent difficile à respecter. Cette difficulté est particulièrement criante lorsque l'essai s'applique à la comparaison des psychothérapies et des chimiothérapies. D'abord, les patients manifestent souvent une préférence, soit pour le recours médical, soit pour la psychothérapie, ce qui peut les conduire à refuser leur affectation si elle ne correspond à leurs desiderata. Ensuite, les personnes intégrées au groupe placebo nécessitent parfois d'être prescrites, notamment par prévention d'un suicide. Ainsi, l'effet exercé par la procédure d'évaluation, par les interactions qu'elle fait naître entre examinateurs et sujets, a « sa part » dans l'évolution de l'affection, quand la procédure ne doit pas être adaptée tout bonnement en vertu des symptômes exprimés par le patient.

5) Le dernier problème soulevé par Garfield a trait à l'appréciation des résultats de l'évaluation. D'abord, on a pu observer que le jugement des patients à propos de leur amélioration était très faiblement corrélé avec celui des cliniciens, ($r=0.21$)⁴¹. Selon Garfield, cette divergence résulte des termes mêmes d'une procédure qui exige des thérapeutes de se prononcer sur leur propre pratique. Dans ce domaine, l'introduction d'un superviseur ou d'un juge indépendant ne résout qu'imparfaitement la question : les notations de ces évaluateurs se rapprochent de celles des patients mais restent toujours inférieures à celles des thérapeutes. Quelle évaluation retenir ? Jusqu'à ce jour, aucune réponse n'a été apportée à ce problème. La durée à réserver au traitement fait, de la même manière, l'objet d'un débat qui n'a pas été tranché. Si une thérapie n'apporte pas d'amélioration significative au bout de 6 mois, doit-on pour autant conclure à son absence d'efficacité ? Plus globalement, la question sur laquelle bute l'évaluation chiffrée du rendement thérapeutique, concerne la délicate conversion de l'information statistique au propos de la clinique. Pour une échelle en 13 points, par exemple, toute notation inférieure à 6 à l'issue de l'essai pourra être associée à une amélioration significative. Sur un plan clinique, cela signifie qu'une symptomatologie résiduelle est encore présente. Dans pareil cas, faut-il envisager

⁴⁰ M. Zimmerman M., J.L. Mattia, M.A. Posternal, "Are subjects in pharmacological trials of depression representative of patients in routine clinical practice?", *American Journal of Psychiatry*, 159 (3), 2002, pp. 469-473. Les auteurs rapportent que parmi 346 patients souffrant de dépression majeure, traités en ambulatoire, seuls 41 d'entre eux (soit 12%) auraient pu participer à un essai clinique. L'exclusion probable de la majorité des patients obéit au fait que la plupart d'entre eux présentent une co-morbidité (syndrome bipolaire, dépression psychotique, anxiété, addiction, idées suicidaires) ou une intensité dépressive légèrement inférieure au seuil d'inclusion.

⁴¹ R. B. Sloane, F.R. Staples, A.H. Cristol, N.J. Yorkston, K. Whipple, *Psychotherapy versus Behavior Therapy*, Cambridge, Harvard University Press, 1975.

de poursuivre le traitement ? Cette comptabilité s'accorde sur un seuil qui enferme en réalité une multitude de cas cliniques : la note représentant la somme d'une série indifférenciée de symptômes, on trouvera, parmi les patients dont l'état s'est amélioré, des personnes dont l'humeur, l'idéation ou bien encore le rythme psychomoteur reste dégradé, c'est-à-dire autant de signes réclamant une prise en charge spécifique.

L'article de Garfield se conclut sur l'idée suivante : plus on formalise le protocole des essais cliniques, plus on s'éloigne de la pratique réelle des cliniciens⁴². L'argument, si on le déplie jusqu'à son terme, entend indiquer que les conditions d'interaction, de décision et d'interprétation imposées par l'évaluation standardisée sont sans commune mesure avec celles qu'exige la relation clinique.

Une efficacité trouble

Sur quels résultats débouchent ces évaluations ? Et quelles conclusions en sont déduites ? L'article de Lambert et Bergin, publié dans le même ouvrage que celui de Garfield, apporte quelques réponses à ces questions⁴³. Selon les auteurs, 80% des patients engagés dans une psychothérapie connaissent une amélioration significative. En revanche, en l'absence de traitement, le taux de rémission ne serait plus que de 40%. L'efficacité des psychothérapies est patente pour l'ensemble des modèles étudiés, lesquels sont d'ailleurs pour la plupart de facture proprement anglo-américaine (TCC, psychanalyse dynamique, thérapie interpersonnelle). Selon un fait souvent rapporté par la littérature, les thérapies comportementales donneraient lieu à de meilleurs résultats dans le traitement des phobies, tandis que les effets des thérapies cognitives seraient plus manifestes dans la prise en charge de la dépression. L'efficacité des psychothérapies serait plus manifeste pour les pathologies les moins sévères, et plus résiduelle dans le traitement de la schizophrénie, des addictions ou des désordres de la personnalité. Toutefois, selon Lambert et Bergin, ces limites n'enlèvent rien à la performance des modèles psychothérapiques : leur action est significative, en dépit de l'évidente diversité des approches mobilisées⁴⁴.

Selon les auteurs, que l'efficacité des psychothérapies puisse ainsi se jouer des clivages d'obédiences et de pratiques oblige à conclure à l'existence d'un principe d'action partagé par l'ensemble des modèles. Par extrapolation, le constat laisse à penser que la psychothérapie, quelle qu'en soit la forme, est douée d'une qualité thérapeutique qui lui est spécifique. Laquelle ? Lambert et Bergin font ici valoir une première hypothèse, souvent mentionnée par l'évaluation : la

⁴² « Plus l'essai expérimental est contrôlé, plus les cliniciens sont dûment sélectionnés et entraînés, plus les patients sont déterminés à participer à la recherche et sont susceptibles de demeurer dans le programme, plus les mesures établies avant et après la cure sont formalisées, plus les critères de sélection des patients sont rigoureux, plus sera grand l'écart entre les résultats de la recherche et ceux de la pratique », S.L. Garfield, "Major issues in psychotherapy research", *art. cit.*, p. 337.

⁴³ M. Lambert, A. Bergin, "Achievement and limitations of psychotherapy research", in D. Freedheim (ed), *History of Psychotherapy. A Century of Change*, *op. cit.*, pp. 360-390.

⁴⁴ A vrai dire, ces procédures évaluent les modèles les mieux représentés au sein de l'enseignement et de la recherche : autant dire, qu'on ne dispose à travers elles d'aucune information tangible sur les performances des quelques 400 techniques psychothérapeutiques recensées pour les Etats-Unis.

psychothérapie agirait comme un placebo. Sous ce jour, elle pourrait être comparée à un pseudo-traitement qui redonne espoir au patient, simplement en vertu de l'attention qui lui est prêtée à cette occasion. On est perplexe. Les essais pharmaceutiques entendent distinguer la part respective des effets physiologiques et psychologiques, en vue d'établir le caractère proprement actif du médicament testé. Dans ce cadre, dire que la substance évaluée agit comme un placebo, c'est considérer qu'elle est dénuée en propre d'une propriété thérapeutique. Aussi, comment les psychothérapies pourraient-elles revendiquer une action et la qualifier, à la fois, d'effet placebo ?

La revue de littérature conduite par Shapiro en 1982 a permis d'ouvrir une autre perspective, sans doute moins problématique⁴⁵. A partir de la recension des études consacrées au traitement de la dépression, l'auteur parvient à la conclusion suivante : si les thérapies par antidépresseurs obtiennent les résultats les plus probants, les taux d'amélioration imputables aux psychothérapies n'en restent pas moins supérieurs à ceux obtenus au moyen d'un médicament placebo. L'étude de Shapiro suggérerait donc que les techniques psychothérapeutiques étaient bel et bien douées d'une propriété thérapeutique. Mais, si celle-ci ne se réduit pas à un pur effet de suggestion, quelle en est alors la nature ?

Selon Lambert et Bergin, la littérature de l'évaluation aurait mis au jour deux mécanismes. Le premier a trait à ce qu'on pourrait appeler « l'effet thérapeute ». La sensibilité, l'honnêteté, la disponibilité d'écoute du clinicien conditionneraient pour beaucoup l'issue de la thérapie. D'après les auteurs, plus de 70% des patients traités avec succès accordent d'ailleurs à cette dimension un rôle cardinal⁴⁶. Logiquement, le second facteur que Lambert et Bergin mettent en avant a trait à « l'effet patient ». Ainsi, Strupp, comparant l'action engagée par des thérapeutes auprès de groupes dépressifs et anxieux, fait remarquer que la probabilité de l'amélioration, au-delà de l'affection considérée, varie de manière significative selon que le patient coopère ou démontre au contraire une attitude de désengagement, voire d'opposition⁴⁷. C'est pourquoi, selon Strupp, la personnalité, l'organisation de l'ego, la maturité, la motivation et la capacité à s'engager de manière productive dans une thérapie verbale représentent, du côté du patient, autant de dispositions susceptibles d'influer sur le cours du traitement.

De ces obstacles à la thérapie, se déduisent par contraste les mérites qui lui sont attribués : l'action de la psychothérapie obéit, selon ces évaluateurs, à la qualité de la relation qu'elle permet d'instaurer. L'argument n'est pas sans faire écho à la particularité des thérapies américaines. En effet, celles-ci s'informent à une tradition qui met avant l'interaction du thérapeute et du patient et l'égalité de leur échange, à la différence du modèle freudien qui prône le maintien d'une stricte asymétrie de relation entre les deux acteurs de la cure. Il est vrai que, pour les évaluateurs, cette divergence de perspective compte peu puisque, selon eux, les particularités des techniques déployées ne suffisent pas à expliquer l'action des psychothérapies. A la dimension technique, il conviendrait d'ajouter un facteur

⁴⁵ D. A. Shapiro, K. Shapiro, "Meta-analysis of comparative outcome studies: A replication and refinement", *Psychological Bulletin*, 92, 1982, pp. 581-564.

⁴⁶ Plus exactement les qualités exigées semble-t-il du thérapeute sont les suivantes : soutien, chaleur, empathie, réassurance, suggestion, crédibilité et encouragement.

⁴⁷ H.H. Strupp, "Success and failure in time-limited psychotherapy. A systematic comparison of two cases—comparison. 4, *Archives of General Psychiatry*, 37, 1980, pp. 947-954.

relationnel, jugé plus déterminant. En réalité, la qualité de ce facteur est pour le moins ambiguë puisque la relation est considérée ici tout à la fois comme la condition, la propriété et l'objectif de la thérapie. Surtout, il y a là un truisme, les psychothérapies procédant, par définition, d'une technique relationnelle.

CONCLUSION

La psychopathologie expérimentale plaçait dans la clinique la possibilité de fonder en preuves les théories explicatives de l'affection mentale. Elle y cherchait un terrain propice à l'observation des phénomènes psychiques, ou bien encore un moyen d'affiner son action sur la pathologie. Avec l'évaluation standardisée, la psychopathologie s'est ouverte à la biomédecine, à la statistique des essais randomisés et à une certaine forme de l'expérimentalisme médical, fondé sur l'étude des conséquences de l'intervention thérapeutique. Aux Etats-Unis, dans un contexte mettant en jeu l'utilité sociale de la psychiatrie, de la psychologie clinique et de leurs thérapies, la psychopathologie quantitative a été ainsi portée à inscrire le diagnostic et le travail thérapeutique au centre de son évaluation et a été, finalement, promulguée en arbitre. La clinique était ainsi faite objet et formellement distinguée de l'espace de recherche. Ce changement fut prétexte à la réorganisation des rivalités qui, depuis la naissance même des études sur la folie, opposent les tenants d'une lecture biogénétique aux partisans d'une approche psychogénétique de ce phénomène. A ce jeu, comme on le sait, la psychopharmacologie a pris l'ascendant sur les psychothérapies, et l'évaluation standardisée l'a emporté sur la clinique psychanalytique, du moins, en Amérique.

Mais, l'évaluation a également ouvert des querelles inédites entre nosographes et psychométriciens, psychiatres et psychologues. La transposition de la méthodologie des essais cliniques à la psychopathologie a commandé en effet de trancher une série de questions. L'organisation de l'évaluation et de ses protocoles, la représentation même de la pathologie ont été débattues, et ont pu finalement bénéficier d'un règlement technique, fondé sur une répartition bilatérale des tâches et des mesures entre les spécialités. Ce règlement a permis de résoudre, par le jeu des outils et des procédures, l'épineux problème de la caractérisation du changement en psychopathologie, problème qui ne pouvait être abordé sur le fond ni par l'approche par syndrome de la tradition classificatoire ni par la théorie des dispositions de la psychométrie.

Mais, c'est sans nul doute à travers l'extension de ce programme aux psychothérapies que l'évaluation clinique a atteint, d'une certaine manière, son point limite. Ici, l'amendement à la méthodologie des essais a été plus que sérieux puisque le recours au double aveugle et au placebo ne pouvait y être pratiqué. Surtout, il y a lieu de s'interroger sur l'intérêt de cette transposition de méthode. Le dispositif conçu par la biomédecine est né de l'intention d'isoler, pour une maladie donnée, la part proprement active des thérapeutiques en vue d'en spécifier l'intérêt médical. A cette fin, les essais cherchent à rendre inopérantes les variables susceptibles d'interférer avec la mesure des performances intrinsèques de la thérapeutique, à commencer par l'interaction du clinicien et du patient. Si on admet que l'action de la psychothérapie procède d'un travail de la relation, lui appliquer le modèle des essais revient donc, en substance, à vouloir déterminer

son rendement au moyen d'une méthode qui est censé en neutraliser l'impact. En dépit d'une telle gageure, l'évaluation est parvenue à montrer que les psychothérapies possèdent une part significative d'efficacité, supérieure à celle d'un placebo. Selon les évaluateurs, leurs propriétés étaient donc patentes. Pour qualifier cette action, l'évaluation a distingué la part d'influence exercée par chacune des techniques psychothérapeutiques testées, qu'elle a jugée insuffisamment discriminante, d'une propriété plus générique et plus déterminante, qu'elle a imputée aux bienfaits de la relation thérapeutique elle-même. C'est là, plus qu'un paradoxe, une contradiction, puisque la dimension relationnelle, proprement exclue par le protocole d'évaluation, était à cette occasion de nouveau tenue pour explicative.

La mesure du rendement psychothérapeutique donnerait sans doute lieu à des résultats plus riches si elle n'excluait pas, par principe, le recours à une théorie explicative du trouble mental et des mécanismes de la relation thérapeutique. Mais, cette perspective reviendrait à renouer avec une méthode fondée sur la mise à l'épreuve d'hypothèses, c'est-à-dire avec le programme auquel l'évaluation a cherché précisément à s'opposer, en se réclamant d'une orientation pragmatique et agnostique. Il y a là une manière de comprendre la saisissante diffusion de la démarche évaluative puisque, avec elle, la recherche paraît disposer d'un modèle d'investigation capable de transcender toutes les rivalités de points de vue.

Une autre manière d'enrichir la recherche consisterait à l'appliquer à des modes de prises en charge, certes préalablement typés, mais réels (hospitaliers, ambulatoires, sectoriels, etc.). Autrement dit, l'étude de la clinique gagnerait à partir de la clinique elle-même, plutôt que des dispositifs d'interaction composés par la statistique expérimentale. Mais, il y a peu de chances que cette option soit retenue. En effet, il y a entre la proposition d'une recherche en situation et la démarche de l'évaluation standardisée, plus qu'un abîme, un désaccord de fond sur la vision de l'objet, des données et de l'expérimentation. Maurice Halbwachs avait distingué, il y a déjà longtemps, deux philosophies radicalement différentes de la mesure, l'une qu'il destinait aux objets physiques, l'autre, aux objets sociaux. On ne saurait mieux résumer la distance qui sépare la démarche évaluative d'une statistique appliquée à l'analyse des relations humaines :

« Une série de mesures peut être utilisée de deux manières bien différentes. Ou bien on se propose de connaître une grandeur qui a une existence au moins apparente hors de nous. Alors, celui qui mesure devra s'efforcer de se subordonner à ce qu'il perçoit, et d'échapper, tandis qu'il observe, à toutes les influences que les membres de son groupe ont pu ou peuvent exercer sur lui. [...] Ou bien on s'intéresse aux observations, on cherche à déterminer [...] la façon dont les hommes de différents groupes (groupes réels et non constitués au hasard), perçoivent et mesurent un même objet apparent. Mais, alors; il ne faut plus que l'individu se subordonne étroitement à la grandeur, jusqu'à sortir de son groupe et à s'oublier lui-même : il vaut même mieux que l'objet soit tel qu'il évoque, dans l'esprit de l'observateur, des représentations familières, et qu'il lui donne occasion de commettre le genre d'erreurs qui sont propres à la société dont il est membre [...] »⁴⁸,

⁴⁸ M. Halbwachs, « L'expérimentation statistique et les probabilités », *art. cit.*, p. 362.