



HAL
open science

Un conflit entre la méthode et son objet : l'application des essais cliniques à l'évaluation des psychothérapies

Philippe Le Moigne

► To cite this version:

Philippe Le Moigne. Un conflit entre la méthode et son objet : l'application des essais cliniques à l'évaluation des psychothérapies. Tarik Benmarhnia; Pierre-Marie David; Baptiste Godrie. Les Sociétés de l'expérimentation. Enjeux épistémologiques, éthiques et politiques, Presses de l'Université du Québec, pp.81-102, 2019, Problèmes sociaux et interventions sociales, 978-2-7605-5109-1. hal-03434751v2

HAL Id: hal-03434751

<https://hal.science/hal-03434751v2>

Submitted on 30 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.

D'un conflit entre la méthode et son objet

L'application des essais cliniques à l'évaluation des psychothérapies

Philippe Le Moigne

Sociologue, INSERM – CERMES3, 45 rue des Saints-Pères, 75006 Paris

philippe.le-moigne@paris5.sorbonne.fr

Les origines, tant médicales, psychologiques que statistiques de l'essai clinique ont structuré sa forme canonique sous le mode dit « randomisé, avec placebo, en double aveugle ». Le fait d'interroger le schéma d'organisation de cette méthode, par un bref retour sur son histoire, permet d'en situer plus précisément l'objectif et autorise, ensuite, à mieux appréhender l'intérêt de sa diffusion actuelle. En effet, la méthode sert aujourd'hui à mesurer aussi bien l'impact des politiques du développement qu'à évaluer l'efficacité des programmes de nutrition. Autrement dit, l'essai clinique a cessé de se limiter à l'analyse des effets du médicament pour s'appliquer à l'étude des comportements sociaux, sans que la généralisation de cette méthode n'ait toujours été réellement interrogée.

L'application des essais à l'évaluation des psychothérapies permet d'éclairer au plus près certaines des limites de cette transposition. Les essais sont en effet fondés sur deux grands axiomes : la neutralisation des variables d'intérêt marginal (ou « parasites »), à laquelle la randomisation est censée répondre ; le contrôle de l'interaction entre médecins et patients qui vise en particulier à limiter les effets de suggestion induits par l'activité de prescription : d'où les fonctions prêtées au placebo et au double aveugle. En revenant sur le développement de ce mode de raisonnement dans le champ des psychothérapies, il devient possible d'observer la part des difficultés sur lesquelles cette transposition de méthode est susceptible de buter. Cette étude est instructive dans la mesure où elle illustre le cas, précisément, d'un conflit de « nature » entre la méthodologie des essais et l'objet qu'elle est censée évaluer¹.

Introduction

Le recours aux essais cliniques fait depuis quelques années florès, au point d'avoir pour large part quitté son terrain de prédilection, celui de l'efficacité des médicaments (Bourret et Le Moigne, 2014). La méthode sert aujourd'hui à aborder des questions et des objets dont la nature ne se prête pas immédiatement aux cadrages que l'essai clinique impose,

¹ Je tiens à remercier Tarik Benmarhnia et Baptiste Godrie qui, par leurs critiques judicieuses, ont beaucoup œuvré à l'amélioration de ce texte.

sans que ce déplacement n'ait toujours été interrogé. Déjà, dans le champ d'étude des produits de santé, tels que les vaccins, l'introduction de l'essai clinique bute sur des logiques sociales qu'il n'est pas toujours en mesure de contrôler, comme le démontre à l'envie l'expérience de la recherche sur le VIH en Afrique. Ici, les conditions imposées par la procédure, telles que l'observance thérapeutique et l'assiduité aux visites médicales, ne peuvent être réunies que dans la mesure où l'espace créé par la recherche réussit à mobiliser patients et médecins à sa cause. La perspective implique ainsi que la médecine traditionnelle soit mise à distance, et que les routines de la consultation médicale soient réinvesties et réorientées par les praticiens locaux en vue de servir l'enjeu de la recherche et seulement lui. Faute d'y parvenir toujours, l'essai clinique est tenu de déployer de multiples ajustements, en s'attelant par exemple à la correction rétrospective des erreurs d'enregistrement, ou bien encore en mobilisant médecins, infirmières et aides-soignantes dans le but de rechercher et de réintégrer les « perdus de vue » dans le protocole (Brives, 2015).

Cet exemple suffit à montrer que *la mise en place d'un essai clinique ne se conçoit pas sans le préalable d'un contrôle relationnel* : le protocole doit avoir acquis, au moins pour partie, la maîtrise des échanges indispensables à l'inspection du fait pour laquelle il a été conçu. Or, la question de ce contrôle comprend une difficulté supplémentaire lorsque l'objet étudié tend à s'appliquer, comme aujourd'hui, à des comportements de santé ou bien à la pratique médicale elle-même, c'est-à-dire à des phénomènes dont l'économie est proprement relationnelle. Par exemple, certaines études cherchent aujourd'hui à mesurer l'amélioration des comportements alimentaires de la population au moyen d'une procédure d'éducation à la santé, en l'occurrence, par la délivrance durant une période donnée d'un panier quotidien de fruits et de légumes à un groupe dit « traité », le groupe témoin en étant privé (Campbell *et al.*, 2013). Si tant est qu'on puisse mesurer ainsi un changement de comportement, à quoi l'assigner ? À l'efficacité du dispositif éducatif ou bien au fait que le protocole ait offert un bénéfice net à cette population en termes de provision de nourritures ? Et, en l'absence de suivi, comment être certain que les changements alimentaires observés ne cesseront pas lorsque ce bénéfice aura disparu ? En s'appliquant ainsi aux comportements de santé, le contrôle relationnel requis par l'essai peut tout simplement déboucher sur un artefact : le changement escompté, étant seulement produit par la procédure, disparaît lorsque celle-ci est éteinte.

On pourrait appliquer les mêmes réserves aux essais qui s'appliquent à dégager les contours des « bonnes pratiques » dans le champ médical (Bourret et Le Moigne, 2014). Ce type de protocole cherche à mesurer méthodiquement les avantages d'un procédé technique,

d'une démarche diagnostique ou encore d'une procédure de prise en charge. Que ce genre d'études puisse démontrer qu'une pratique de soin est plus efficace qu'une autre s'envisage aisément. Mais la question reste de savoir si les conditions dans lesquelles cette efficacité a été mesurée ont à voir, de près ou de loin, avec les conditions réelles de l'exercice médical. Autrement dit, quand bien même les praticiens tireraient bénéfice sur un plan médical des pratiques ainsi préconisées par la recherche, il n'est pas certain qu'ils puissent ou cherchent à les mettre en œuvre par la suite dans le contexte de leur exercice professionnel. De fait, cet exercice dépend d'interactions avec la hiérarchie, les confrères ou encore avec les patients, c'est-à-dire de contraintes et d'enjeux relationnels que la forme de l'essai clinique a exclus pour ne considérer en propre que l'efficacité de la technique ou du mode d'intervention visé par la recherche.

Ce texte voudrait poursuivre la réflexion engagée à la croisée des deux problématiques d'action, brièvement pointées ici, en mêlant la question du *contrôle de l'interaction* à celle de *la qualité de l'objet* à laquelle s'adresse l'essai clinique. Sur ce dernier point, il n'est pas simplement question de faire remarquer que la maîtrise acquise par l'essai clinique sur l'interaction sociale varie en fonction des pathologies, des comportements étudiés, et des enjeux que leur traitement ou que leur évaluation soulève. *Le propos vise plutôt à faire remarquer que cette maîtrise peut conduire, dans certains cas, à l'imposition d'un « effet de procédure », l'essai recouvrant de ses principes le phénomène qu'il est censé étudier au point de le transformer à son image.* Naturellement, ce renversement est d'autant plus probable que l'objet d'étude s'applique aux comportements de santé et, plus encore, à l'interaction thérapeutique elle-même.

Pour donner à voir cette éventualité, l'application des essais cliniques à l'évaluation des psychothérapies constitue un terrain de choix. Il ne s'agit pas de revenir sur les controverses suscitées, en France, par l'expertise de cette évaluation par l'Institut National de la Santé et de la Recherche Médicale – INSERM (Le Moigne, 2008). Il s'agit plutôt d'observer comment l'application de l'essai clinique, en vertu même de ses principes, bute ici sur un point-limite. Pour alimenter cette réflexion, le texte revient d'abord sur la genèse des principes de l'essai clinique, puis sur leur transposition dans le cadre de la recherche consacrée aux psychothérapies et, enfin, sur les enseignements de cette recherche².

² Il y aurait lieu de comparer les difficultés soulevées par l'emploi des essais cliniques dans l'évaluation des psychothérapies avec celles qui ont pu être notées pour d'autres terrains, notamment dans le domaine de l'économie du développement (Bédécarrats *et al.*, 2013). Faute de place, on se contentera de signaler les pistes de réflexion qui ont été engagées autour d'une interrogation plus générale sur les conditions d'« applicabilité » de la méthode (Treweek et Zwarenstein, 2009).

1. Retour sur la méthode : une prédilection pour le hasard et l'ignorance

Dans leur définition la plus simple, 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 une thérapie de substitution. La distribution 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'État 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. De fait, l'idée d'évaluer l'effet d'une mesure publique de façon purement « étanche » pouvait sembler illusoire, compte tenu de la multiplicité des facteurs agissant sur la vie sociale. 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, simplement parce qu'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 ». Pour les sociologues français de l'École durkheimienne en particulier, 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 sociale³.

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 (McCall, 1923). Dans les années 1930, le statisticien Ronald Fisher conclut 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 différences observées entre les groupes peuvent être imputées au seul fait d'avoir reçu le traitement, il élabore un test statistique resté fameux : l'analyse de co-variance (Fisher, 1935). 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 ?

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

³ On trouve chez Durkheim l'illustration directe de cet 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 (Durkheim, 1895). Maurice Halbwachs développera dans les années 1920, le même argument (Halbwachs, 1923). Ce n'est que dans les années 1930 que la sociologie américaine tentera, à travers le travail de Francis Chapin en particulier, d'appliquer l'expérimentalisme aux faits sociaux (Chapin, 1947), sans toutefois qu'une telle transposition n'ait réussi à perdurer au-delà des années 1980 (Oakley, 1998).

porteuses d'amélioration, en vertu d'effets dits « psychologiques », assez mal identifiés⁴. 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 suggestion (Kaptchuk, 1998).

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 bien (Marks, 1999). À partir des années 1950, l'expérimentalisme médical 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 aveugle⁵.

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 ? Si on se réfère à Ulrich Tröhler, le projet du double aveugle naît également au 18^{ème} siècle (Tröhler, 2000). 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

⁴ 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 attribuée 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 (Kirsh, 1997)

⁵ Pour une présentation canonique de la procédure, se reporter à Pocock (1983).

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 nouveau (Kaptchuk, 1998). 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.

2. L'application des essais aux psychothérapies : un ajustement contraint

La psychopathologie quantitative, nom aujourd'hui donné à la psychiatrie fondée sur la mesure, a repris à son compte les termes du dispositif imaginé par la statistique et l'experimentalisme médical (Guelfi *et al.*, 1995). Elle en a fait sa raison, il est vrai, non sans avoir procédé à certains ajustements. Mais, avant d'évoquer la teneur de ces aménagements, une question qui reste souvent inaperçue mérite d'être abordée : comment les auteurs des essais cliniques ont-ils conçu d'apprécier l'amélioration des patients atteints d'un trouble psychique, 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 ?

Premier dilemme : mesurer l'efficacité de la clinique mentale

Pour la psychiatrie, le cadre logique de l'essai clinique marque un resserrement de l'analyse sur l'évolution de la pathologie, au détriment de sa compréhension. 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 États-Unis, le *Psychopharmacological Service Center*, qui préfigure la naissance de la psychopharmacologie, se rallie à ce découpage de la preuve dès 1957. Outre-Atlantique, le Kefauver-Harris Amendment de 1962 (Healy, 1997), qui fait suite à la mise en évidence des effets tératogènes de la thalidomide (preuve ayant été faite de malformations chez les enfants nés de mères ayant consommé ce produit), donnera à cette division du travail le caractère

d'une exigence légale : les laboratoires pharmaceutiques doivent dès lors assurer l'État 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 Unity* du *National Institute for Mental Health* (Guy, 1976). Cette durée de gestation pourrait être imputée aux difficultés que 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.

1) Au début des années 1950, R.D. Wittenborn, produit l'une des premières grilles d'évaluation dédiées à la psychopathologie (Wittenborn, 1955). Ce psychologue psychométricien cherche à produire 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 des états de rémission ou d'aggravation de la maladie. 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. Évidemment, 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 Scale* (Overall et Gorham, 1962). Il est directeur en chef du Bureau des Standards de la très influente *Veteran Administration* qui veille à la prise en charge 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 au sein de la population à la manière d'une distribution normale (Spearman, 1927). 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'étagier la pathologie mentale, depuis ses occurrences les plus frustes jusqu'à ses formes les plus sévères, sur un simple continuum. Suivant cette hypothèse, les névroses étaient placées sur le gradient le plus proche de la normalité, tandis que les psychoses étaient positionnées au sommet de l'échelle pour leur gravité. Ainsi bâti, son instrument pouvait en théorie 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, fondée sur le répertoire diagnostique des maladies. Cette tradition, dominée par la classification des maladies de Kraepelin, est catégorialiste : les entités cliniques y sont comprises comme des classes distinctes. Elles relatent le fait que, *chez certains groupes d'individus seulement*, une série particulière de symptômes est réunie, à la manière de *clusters*, avec une fréquence qui ne peut être imputée au hasard. Dans la mesure où c'est ici la spécificité des agrégats de symptômes (ou syndromes) 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 schizophrénies. Celui-ci, comme le diagnostic de l'ensemble des troubles mentaux, demeurera sous le monopôle de la psychiatrie et de la tradition classificatoire.

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 remplis par le patient, alors que les échelles diagnostiques étaient renseignées par le praticien au décours de la consultation. La généralisation de la procédure par questionnaire à l'évaluation de la schizophrénie était difficile, compte tenu de l'incapacité des patients. 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 arbitrage 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. 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* (Cattell et Kline, 1977). Mais ils ne peuvent être précisément appréhendés par la discipline dans la mesure où la psychométrie n'envisage pour objet que des structures stables, le plus souvent héritées, et donc peu sujettes à variations.

Pour tenter d'apporter un règlement à la question du changement, l'évaluation a cherché à lier plus ou moins explicitement les approches catégorialistes 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 : à l'instar du *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III), publié en 1980 (APA, 1980), ces échelles présentent, pour chacun des troubles recensés par la nosographie, la liste des symptômes requis pour qu'un diagnostic positif puisse être retenu. De leur côté, les tests des psychologues, reconstruits sous la forme d'échelles se prêtant à l'hétéro-évaluation, à l'instar de l'échelle d'Hamilton pour la dépression (Hamilton, 1960), ont été réorganisés en vue de satisfaire bon an mal an aux exigences d'une mesure du changement. À 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 (ou sévérité). 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 patient. Ce chiffrage est entamé lors de la mise en œuvre 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 clinique⁶.

Ce *modus vivendi* a permis à la fois d'apaiser les rivalités disciplinaires et d'organiser la division du travail entre les différents producteurs d'échelles. Cette organisation instrumentale a, par ailleurs, offert une solution tangible à la caractérisation du changement, particulièrement utile pour jauger l'efficacité des médicaments et des thérapies de la clinique mentale. Toutefois, *il reste que cette évaluation de l'efficacité est d'abord métrique, instrumentale et délocalisée, c'est-à-dire insensible par principe au contexte. Or, toute la question demeure de savoir si, dans le domaine de la clinique mentale, la validation d'un effet tangible de la cure peut se départir d'une estimation interpersonnelle et située, c'est-à-dire d'un accord auquel le clinicien et le patient (ou encore un collègue de praticiens) donnent foi, au décours de la consultation*⁷. Autrement dit, le problème soulevé par *le caractère abstrait* de l'évaluation produite au moyen de ces échelles peut se lire comme suit : il n'y est pas seulement question d'un probable excès de formalisation, mais du fait qu'un tel excès pourrait contribuer à soustraire l'objet à mesurer de la mesure elle-même. Ce risque apparaîtra de manière plus claire, et sera dénoncé comme tel plus franchement encore, lorsqu'il s'agira de soumettre le déroulé des psychothérapies aux règles protocolaires des essais.

Second dilemme : la soumission des psychothérapies aux règles protocolaires des essais

La division du travail d'évaluation entre nosographes et psychométriciens a fortement marquée l'organisation des essais cliniques dédiés aux traitements médicaux du trouble mental. L'essentiel de son esprit a été conservé lorsque l'évaluation s'est appliquée, au-delà des études de la psychopharmacologie, à la mesure d'efficacité des psychothérapies. Deux raisons au moins ont présidé à cette extension de l'évaluation. Il s'est d'abord agi de proposer les psychothérapies au remboursement des soins, sur la base d'une efficacité démontrée (Barton, 1987). Il s'est agi ensuite, non plus de comprendre le fonctionnement de ces

⁶ 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 (Pichot et Olivier-Martin, 1974).

⁷ C'est d'ailleurs pourquoi la psychopathologie quantitative a créé un coefficient statistique inter-juge, dit *kappa*, en vue de valider ces échelles. Dans cette perspective, la plupart des grilles diagnostiques du DSM-III ont été mises à l'épreuve, le coefficient visant à établir si des cliniciens, confrontés aux mêmes patients, parviendraient à renseigner ces échelles avec suffisamment de convergence (Spitzer et Fleiss, 1974). Au-delà des mérites statistiques du coefficient, d'ailleurs largement décriés, il reste qu'un tel projet est pour le moins significatif : en soumettant la validation de leurs instruments à « l'épreuve de la consultation », leurs concepteurs laissaient entendre *in fine* que seule l'interaction clinique pouvait constituer le juge de paix de l'évaluation.

thérapies à proprement parler, mais d'en certifier l'efficacité par un examen statistique⁸. 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 assésés, entre autres, par l'antipsychiatrie et sa dénonciation du régime asilaire et de l'arbitraire diagnostique des praticiens (Grob, 1995). Dans ce cadre, la clinique dite « mentale » 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 nombres.

Ici, toutefois, la rigueur de l'expérimentalisme a dû concéder certains aménagements de procédure, pour au moins trois raisons. 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 confrère⁹.

En second lieu, à la différence des thérapeutiques médicinales qui peuvent circuler sans que leur structure moléculaire n'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 évènement 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

⁸ Paul Meehl, psychanalyste et psychométricien, a formulé l'une des premières défenses argumentées de ce principe (Meehl, 1954).

⁹ Le fait que les thérapeutes soient conduits à juger eux-mêmes des effets de la cure comporte un biais qui a été désigné sous terme d'« effet dodo », en référence à l'oiseau d'*Alice au pays des merveilles* à qui il était demandé de désigner le vainqueur d'une course autour d'un lac et qui, en réponse, déclara que tout le monde avait gagné. Dans le cas qui nous intéresse, les cliniciens ayant tendance à noter plus favorablement les modèles thérapeutiques qui ont leur faveur, il s'en suit que les résultats des essais tendent à dégager des effets somme toute égaux entre les différents types de traitement, comme si « chacun d'eux avait gagné » (Luborsky *et al.*, 2002).

s'est appliquée à l'examen d'une clinique dont on peut douter qu'elle corresponde à l'exercice réel de la psychothérapie.

Enfin, en troisième lieu, 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 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. Dans ce dernier cas, la procédure cherche à apprécier dans quelle mesure la proportion de rémissions naturelles (à l'intérieur du groupe témoin) pourrait s'avérer comparable au taux d'amélioration obtenu en propre au moyen de la psychothérapie (au sein du groupe traité)¹⁰. Pour autant, ce découpage est-il tenable ?

3. Les essais à l'épreuve de la relation thérapeutique

Un artefact de procédure

Une des manières de répondre à cette question consiste à se pencher sur ce qu'en disent les évaluateurs eux-mêmes. Quel est, selon eux, le rendement de l'application des essais cliniques à l'évaluation des 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 Association* en 1992, à l'occasion du centenaire de la naissance des psychothérapies (Garfield 1992). Cet article, bien qu'ancien, pointe l'essentiel des critiques que seront assignées par la suite à l'évaluation des psychothérapies au moyen de l'essai clinique (Wampold *et al.*, 2011 ; Shean, 2014). Son intérêt majeur est de montrer combien il est difficile d'assimiler les essais randomisés à une mesure même approchée de la pratique clinique. Voici en quelques points, les éléments les plus saillants de son argumentaire :

Premièrement, 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

¹⁰ Une autre méthode consiste à employer un groupe contrôle dénommé *Treatment as Usual* (TAU). Dans ce groupe, les sujets ne sont pas intégrés au dispositif expérimental *stricto sensu*, mais peuvent recourir au soin de leur choix dans leur environnement immédiat, « comme ils le feraient d'habitude ». Toutefois, la question reste ici de savoir quelle valeur assigner à la comparaison entre groupe traité et groupe contrôle, dans la mesure où les sujets ayant recours au TAU peuvent aussi bien se voir prescrire des médicaments, fréquenter un thérapeute que ne bénéficier d'aucun traitement (Burns, 2009).

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 behavioristes, qui entend incarner un modèle d'activité thérapeutique, mais dont on est incapable de dire s'il informe de quelque façon que ce soit l'exercice quotidien du thérapeute.

Deuxièmement, les patients sélectionnés au moyen des échelles diagnostiques sont, en raison des critères d'inclusion, peu représentatifs des cas examinés par la clinique (Westen, 2003). 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 peuvent en probabilité intégrer les essais dédiés à la psychopharmacologie de la dépression (Zimmerman *et al.*, 2002). 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 la cure. C'est pourquoi, selon Garfield, il y aurait lieu de se demander ce qu'il adviendrait des mesures d'efficacité si on étendait les essais aux publics les moins enthousiastes à l'égard des psychothérapies.

Troisièmement, la distinction expérimentale des groupes témoins et des groupes contrôles 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. 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 effets respectifs des psychothérapies et des médicaments psychotropes. D'abord, les patients manifestent souvent une préférence, soit pour le recours médicinal, soit pour la psychothérapie, ce qui peut les conduire à refuser leur affectation si elle ne correspond pas à *leurs desiderata*. Ensuite, les personnes intégrées à la file d'attente nécessitent parfois d'être prescrites, notamment par prévention d'un suicide. Ainsi, la procédure d'évaluation, par les interactions qu'elle fait naître entre examinateurs et sujets, peut avoir « sa part » dans l'évolution de l'affection, et réclamer une adaptation en vertu des symptômes exprimés par le patient.

Enfin, le dernier problème soulevé par Garfield a trait à l'appréciation des résultats de l'évaluation. En particulier, la durée à réserver au traitement a fait l'objet d'un débat qui n'a pas non plus 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 ? (Knekt *et al.*, 2008). 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 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 spécifique¹¹.

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 cliniciens. Il écrit ainsi :

« 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 » (Garfield, 1992 : 337).

L'efficacité des psychothérapies : une action trouble

Au-delà de la dimension d'artefact de la procédure, 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 questions (Lambert et Bergin, 1992). 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 (thérapies comportementales et cognitives, psychanalyse dynamique, thérapie interpersonnelle, etc.).

Selon les auteurs, que l'effet 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

¹¹ Le manque d'isomorphie entre les principes de la quantification et ceux de la clinique a conduit à mettre en doute, d'une autre manière, la lisibilité des résultats issus des essais. Ceux-là s'appuient sur l'évolution moyenne des patients des groupes traités et témoins. Or, cette moyenne s'applique à la somme des changements observés dans les deux populations, sans présager de l'évolution propre à chacun des individus soumis au protocole. Or, pour le jugement clinique, au-delà de toute moyenne, c'est d'abord le critère de l'amélioration individuelle qui permet de valider l'efficacité d'une thérapie (Nadeau, 2012).

l'évaluation : la 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 ici perplexe : si on se réfère au langage des essais, dire que le traitement évalué agit comme un placebo, c'est considérer qu'il est dénué en propre d'une propriété thérapeutique.

La revue de littérature conduite par Shapiro en 1982 permet d'ouvrir une autre perspective (Shapiro et Shapiro, 1982). À partir de la recension des études consacrées au traitement de la dépression, Shapiro 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. Cette étude suggère donc que les techniques psychothérapeutiques sont 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 la nature ?

Lambert et Bergin font valoir ici deux nouvelles hypothèses (Lambert et Bergin, 1992). La première a trait à ce qu'on pourrait appeler « l'effet thérapeute ». La sensibilité, l'honnêteté, la disponibilité d'écoute du clinicien conditionneraient ainsi, pour beaucoup, l'issue de la thérapie (Heinonen *et al.*, 2011). Logiquement, la seconde hypothèse 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 d'une 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'opposition (Strupp, 1980). C'est pourquoi 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 (Clarkin, 2012).

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. À la dimension technique de la cure, il conviendrait donc d'ajouter un facteur relationnel, jugé plus déterminant. Il y a là un truisme, les psychothérapies procédant, par définition, d'une *technique relationnelle* (Wampold *et al.*, 2011).

En effet, la cure ne requiert pas seulement du patient qu'il consente à y participer, comme dans le cas d'un essai clinique. Car, *ni le contrat ni le donné ne sont les mêmes ici* : le premier se fonde sur une mutualité d'attentes, le second puise sa matière dans le produit même de l'interaction. Pour le dire autrement, une psychothérapie ne peut valoir comme telle

que si elle repose sur *le préalable d'un accord* entre le patient et son thérapeute sur le déroulement de la cure, désigné généralement sous le terme d' « alliance thérapeutique » (Horvath A., Bedi, 2002 ; Ardito et Rabellino, 2011), que cet accord assigne à chacun des deux protagonistes une position d'égalité ou d'asymétrie comme dans le cas de l'analyse freudienne. C'est cette réciprocité conditionnelle, étrangère au canevas des essais, qui doit permettre en probabilité de libérer *une possibilité d'échange* entre les deux protagonistes, cet échange constituant par suite *le matériau sur lequel s'appuiera la cure*, et partant, *l'efficacité potentielle* de celle-ci.

Conclusion

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 États-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. Toutefois, l'issue d'un tel processus est d'avoir permis le couronnement d'une évaluation pour le moins formaliste, dont on peut aujourd'hui mieux appréhender les excès.

De fait, à travers l'évaluation des psychothérapies, les essais cliniques se sont appliqués à mesurer les effets de la relation thérapeutique, de sorte, qu'à travers cette extension, la méthode a fini par buter sur un de ses points-limites. De fait, la méthodologie des essais a été ici amendée puisque le recours au double aveugle et au placebo ne pouvait y être pratiqué. Mais, plus avant, il y a lieu de s'interroger sur l'intérêt et les limites de cette transposition de méthode. De fait, 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. À 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 en neutralise 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 cherché à distinguer la part d'influence exercée en propre par chacune des techniques psychothérapeutiques testées, mais elle n'est pas parvenue à montrer qu'un modèle thérapeutique pouvait se prévaloir d'un effet supérieur aux autres. Elle en a déduit que les psychothérapies examinées étaient dotées dans l'ensemble 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.

Quels enseignements tirer de cette relative « dérive » méthodologique ? Autrement, dit pourquoi le déplacement opéré ici par l'expérimentalisme conduit, à bien des égards, à dénaturer le phénomène qu'il s'agit d'étudier par son moyen ? Pour schématiser, on doit concevoir que la mise en œuvre des essais cliniques s'apparente à la formation d'une « scène sociale » à part entière, dotée de ses propres règles. En ce sens, la stabilité comme la fiabilité de cette scène dépendent du respect drastique des procédures de recherche. Mais, à la fois, l'intégrité de cet espace est sans cesse menacée par l'inscription, d'une part, des patients dans leurs milieux sociaux « ordinaires », et d'autre part, des cliniciens dans l'exercice « plus quotidien » de la médecine. De ce point de vue, la scène de l'essai ne peut se concevoir sans la production simultanée d'un jeu d'interactions entre l'espace occupé par le protocole clinique et les cadres sociaux qu'il traverse (Goffman, 1991). Selon les situations, ces interactions peuvent être résiduelles comme elles peuvent prendre l'aspect d'un effet d'immixtion ou encore d'intrusion sur les espaces tiers que le déploiement des essais est appelé à mobiliser. En ce sens, la mise en œuvre de l'essai clinique se lirait mieux comme un jeu d'équations entre *règles prescrites* et *règles locales*, pari étant fait de la subordination de celles-ci à celles-là.

Deux implications se déduisent de cette perspective. D'abord, pour ne considérer que le champ médical, elle suppose la soumission *des comportements de santé* des populations (pratiques de soin, usages médicaux, proximité à la médecine, etc.) *au fait de santé* étudié en propre par l'essai (pathologie, médication, méthode d'intervention, etc.). Ensuite, elle implique une sujétion au moins relative de *la pratique clinique*, donc d'une part significative de l'exercice médical, à *la pratique de recherche* et aux contraintes qu'elle introduit dans l'examen du patient et dans la relation du médecin avec celui-ci. Cette sujétion n'est pas

donnée mais dépend du contexte de l'étude. Autrement dit, le degré d'effacement des règles locales devant les règles prescrites dépend à chaque fois de la nature particulière de l'une et l'autre, c'est-à-dire des spécialités médicales et des milieux sociaux traversés, mais tout également de *l'objet travaillé en propre par l'essai et de ses enjeux*. Cette logique de subordination peut s'avérer désastreuse lorsque l'investigation porte en propre sur un objet tout entier travaillé par une dimension relationnelle, à l'instar de l'interaction thérapeutique. C'est précisément le résultat que donne à voir l'application des essais cliniques sur le terrain des psychothérapies. Cette application produit un effet de clôture qui tend non seulement à soumettre la règle locale à la règle prescrite, mais à la priver de son sens et de son intelligibilité, au point de la vider l'univers de la cure de son objet et d'en rendre l'examen impossible.

En ce sens, ce n'est pas un hasard si la recherche sur l'efficacité des psychothérapies a cherché récemment à emprunter d'autres voies, en recentrant son analyse sur les mécanismes et les conditions rendant compte du mode d'action de la cure, en vue d'en comprendre et d'en améliorer le rendement. C'est en particulier tout le propos des études de cas dites « intensives et naturalistes ». Ces études portent sur l'examen de psychothérapies menées en conditions dites « naturelles », c'est-à-dire produites en conditions réelles. Elle sont dites « intensives » en ce sens où la prospection porte sur des durées longues et examine dans le détail « *l'action thérapeutique à partir des caractéristiques du patient, du thérapeute et de leurs interactions au cours de la thérapie* » (Thurin et al., 2014 : 102). Pour une affection donnée, cet examen entend gagner en généralisation par une analyse qui associe et met en relation l'étude de chaque cas avec celle de l'ensemble des cas réunis par le protocole. Après quoi, l'observation cherche à déterminer quelles postures cliniques et quels types d'intervention thérapeutique paraissent le mieux adaptés à l'amélioration des patients. Ce type d'approche peut constituer l'ensemble du *design* de l'étude ou être intégré, sous la forme d'un modèle mixte, aux protocoles des essais cliniques. Dans pareil cas, l'observation des pratiques en condition réelle cherche à offrir un contrôle externe à la validité des résultats issus des essais et doit permettre, par suite, de statuer sur la possibilité de leur transfert dans la pratique (Midgley *et al.*, 2014).

Il est difficile aujourd'hui d'anticiper la diffusion et l'impact futurs de cette nouvelle orientation dans le champ d'étude des psychothérapies. Mais, on peut y voir comme un aveu : à savoir, que la compréhension de cette pratique thérapeutique requiert de réserver à l'interaction qui la fonde toute l'attention qu'elle mérite, sous peine d'en dénaturer totalement le sens.

Références bibliographiques

American Psychiatric Association, 1980, *Diagnostic and Statistical Manual of Mental Disorders*, Washington – *DSM III*, American Psychiatric Press.

Ardito R.B., Rabellino, D., 2011, Therapeutic alliance and outcome of psychotherapy: Historical excursus, measurements, and prospects for Research, *Frontiers in Psychology*, 270, 2, <https://doi.org/10.3389/fpsyg.2011.00270>

Barton W.E., 1987, *The History and Influence of the American Psychiatric Association*, Washington, American Psychiatric Press.

Bédécarrats F., Guérin I., Roubaud F., 2013, L'étalon-or des évaluations randomisées : du discours de la méthode à l'économie politique, *Sociologies Pratiques*, 27(2), 107-122.

Bourret P., Le Moigne P. (2014 (eds)), Essais cliniques, production de la preuve et mutations de la biomédecine, *Sciences Sociales et Santé*, 32 (3), 5-11.

Brives C., 2015, « Pour eux, c'est seulement à propos de la mortalité ». La production des données et ses conséquences dans le déroulement d'un essai clinique, *Sciences Sociales et Santé*, 33-2, 41-64.

Burns T., 2009, End of the road for treatment-as-usual studies?, *British Journal of Psychiatry*, 195(1), 5-6.

Campbell K.J., Lioret S., McNaughton S.A., Crawford D., Salmon J., Ball K., McCallum Z., Gerner B.E., Spence A.C., Cameron A.J., Hnatiuk J.A., Ukoumunne O.C., Gold L., Abbott G., Hesketh K., 2013, A parent-focused intervention to reduce infant obesity risk behaviors: a randomized trial, *Pediatrics*, 131(4), 652-60

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

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

Clarkin J.F., 2012, An integrated approach to psychotherapy techniques for patients with personality disorder, *Journal of Personality Disorders*, 26 (1), 43-62.

Dehue T., 2000, From deception trials to control reagents: The introduction of the control group about a century ago, *American Psychologist*, 55, 264-269.

Durkheim E., 1895, *Les Règles de la méthode sociologique*, Paris, Alcan.

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

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

Goffman E., 1991, *Les cadres de l'expérience*, Paris, Les Editions de Minuit [1974].

Grob G.N., 1995, The paradox of desinstitutionalization, *Social Science and Modern Society*, 32, 5, 51-59.

Guelfi J.D., Gaillac V., Dardennes R. (dir.), 1995, *Psychopathologie quantitative*. Paris, Masson.

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

Halbwachs M. 1923, L'expérimentation statistique et les probabilités, *Revue Philosophique*, 96, 340-371.

Hamilton M. 1960, A rating scale for depression, *Journal of Neurology, Neurosurgery and Psychiatry*, 23, 56-62.

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

Heinonen E., Lindfors O., Laaksonen M.A., Knekt P., 2011, Therapists' professional and personal characteristics as predictors of outcome in short and long-term psychotherapy. *Journal of Affect Disorders* 2012; 138(3): 301–312.

- Horvath A., Bedi R., 2002, The alliance, In Norcross J.C. *Psychotherapy Relationships that Work: Therapist Contributions and Responsiveness to Patients*, New York, Oxford University Press, 37-70.
- Kaptschuk T.J., 1998, Intentional ignorance. A history of blind assessment and placebo controls in medicine, *Bulletin of the History of Medicine*, 72, 389-433.
- Kirsch I., 1997, Specifying non-specifics: Psychological mechanisms of placebo effects, In Harrington A. (ed.), *The Placebo Effect: An Interdisciplinary Exploration*. Cambridge, Harvard University Press, 166-186.
- Knekt P., Lindfors O., Härkänen T., 2008, Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up, *Psychological Medicine*, vol. 38, no. 5, 689-703.
- Lambert M., Bergin A., 1992, Achievement and limitations of psychotherapy research, In Freedheim D. (ed.), *History of Psychotherapy. A Century of Change, op. cit.*, Washington, APA, 360-390.
- Le Moigne P., 2008, L'efficacité thérapeutique : les psychothérapies de l'expérimentalisme à l'évaluation, in Champion F. (dir.), *Psychothérapie et société*, Paris, Armand Colin, 215-235.
- Luborsky L., Rosenthal R., Diguier L., 2002, The dodo bird verdict is alive and well-mostly, *Clinical Psychology Science and Practice*, vol. 9, no. 1, 2-12.
- Marks H., 1999., *La médecine des preuves. Histoire et anthropologie des essais cliniques*, Le Plessis-Robinson, Institut Synthélabo.
- McCall W.A., 1923., *How to Experiment in Education*, New York, Macmillan.
- Meehl P.E., 1954, *Clinical versus Statistical Prediction. A Theoretical Analysis and a Review of the Evidence*, Minneapolis, University of Minnesota Press.
- Midgley N., Ansaldo F., Target M., 2014, The meaningful assessment of therapy outcomes: Incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression, *Psychotherapy*, 51 (1), 128-137.
- Nadau L., 2012, Les études sur l'efficacité de la psychothérapie ou comment la science répond à une question intime, *Psychotropes*, 18/1, 78-88.
- Oakley A., 1998, Experimentation and social interventions: a forgotten but important history, *British Medical Journal*, 317, 1239-1242.
- Overall J.E., Gorham D.R., 1962, The Brief Psychiatric Rating Scale, *Psychological Report*, 10, 799-812.
- Pichot P., Olivier-Martin R. (eds.), 1974, *Psychological Measurements in Psychopharmacology*, Basel, S. Kruger..
- Pocock S.J., 1983, *Clinical Trial. A Practical Approach*, New York, John Wiley.
- Shapiro D.A., Shapiro K., 1982, Meta-analysis of comparative outcome studies: A replication and refinement, *Psychological Bulletin*, 92, 581-564.
- Shean G., (2014), Limitations of randomized control designs in psychotherapy research, *Advances in Psychiatry*, <http://dx.doi.org/10.1155/2014/561452>
- Spearman C., 1927., *The Abilities of Man. Their Nature and Their Measurement*, New York, Macmillan.
- Spitzer R.L., Fleiss J.L., 1974, A re-analysis of the reliability of psychiatric diagnosis, *British Journal of Psychiatry*, 125, 341-347
- Strupp H.H., 1980, Success and failure in time-limited psychotherapy. A systematic comparison of two cases-comparison, *Archives of General Psychiatry*, 37, 4, 947-954.
- Thurin J.-M., Thurin M., Cohen D., Falissard B., 2014, Approches psychothérapeutiques de l'autisme. Résultats préliminaires à partir de 50 études intensives de cas, *Neuropsychiatrie de l'Enfance et de l'Adolescence*, 62, 102-118.
- Treweek S., Zwarenstein M., 2009, Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials* 2009; 10: 37 (doi: 10.1186/1745-6215-10-37).
- Tröhler U., 2000, *To Improve the Evidence of Medicine. The 18th Century British Origins of a Critical Approach*, Edinburgh, Royal College of Physicians.
- Wampold B, Hollon S, Hill C., (2011), Unresolved questions and future directions in psychotherapy research, In Gelso C.J. (ed.) *The Real Relationship in Psychotherapy: The Hidden Foundation of Change*. Washington, DC, American Psychological Association, 333-355.

Westen D., 2003, Are research patients and clinical trials representative of clinical practice?, In Norcross J.C., Beutler L.E., Levant R.F.(eds.), *Evidence-Based Practices in Mental Health: Debate and Dialogue on the Fundamental Questions*. Washington DC, American Psychological Association, 161-171.

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

Zimmerman M., Mattia, J.L., Posternal M.A., 2002, Are subjects in pharmacological trials of depression representative of patients in routine clinical practice?, *American Journal of Psychiatry*, 159 (3), 469-473.