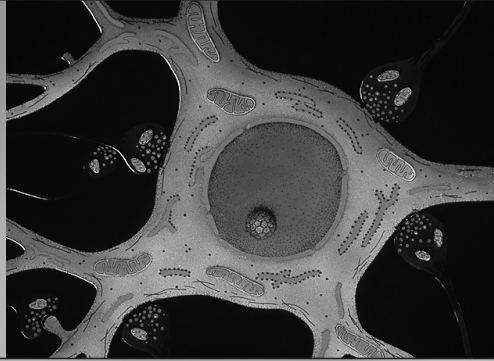
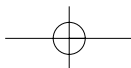
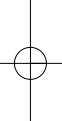
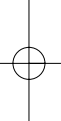
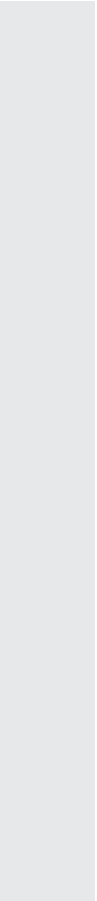
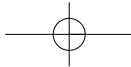


Supplément à Neurone 2003; Vol 8 (N°9)



# Extrapolation des résultats d'études cliniques à la pratique clinique

V Dubois  
A Masson  
B Gillain  
E Stillemans  
B Delatte  
MA Domken  
S Gallemaers  
I Meire



# Introduction

Etant donné que le nombre des nouveaux médicaments augmente, des études comparatives de ces médicaments sont nécessaires pour comprendre leurs points forts et leurs points faibles relatifs et pour guider les décisions en matière de traitement. Lorsqu'elles sont effectuées de manière compétente et analysées correctement, les études cliniques procurent des informations précieuses qui seront interprétées par les médecins en fonction de leur pratique clinique.

Cet article démontre à quel point il est important de comprendre la méthodologie d'une étude clinique car des différences de méthodologie peuvent rendre des comparaisons contestables. Se familiariser avec les différents schémas des études de recherche permet de comprendre et d'interpréter les résultats des études cliniques. Nous aborderons certaines questions qui influencent les données des études cliniques car celles-ci sont extrêmement importantes et informatives pour l'extrapolation des résultats. De plus, il est évident que les résultats probants d'études ayant différents schémas expérimentaux rendent les conclusions plus généralisables.

Les antipsychotiques atypiques constituent actuellement le traitement pharmacologique de première intention recommandé pour les patients atteints de schizophrénie mais on ne dispose que de peu de données pour sélectionner un agent atypique spécifique. A partir de la lecture des études disponibles sur les antipsychotiques (anciens et nouveaux), nous expliquerons comment interpréter les résultats et comment les extrapoler à la pratique clinique.

**L'interprétation des résultats des études cliniques dépend de notre compréhension de la méthodologie utilisée.**

## Qu'est-ce qu'une étude clinique?

Il y a peu de temps encore, le terme d'étude clinique faisait référence avant tout à l'étude clinique randomisée, ou RCT pour "*randomized controlled trial*", qui était considérée comme la référence unique. De nos jours, on appelle étude clinique tout essai structuré qui va mesurer l'effet d'un ou de plusieurs traitements chez des patients. Les caractéristiques de ces essais sont de plus en plus variées, entraînant d'office des possibilités d'extrapolations de différents niveaux. Dans ce paragraphe, nous décrivons différents aspects importants lors de la conduite et de l'interprétation d'une étude. Pour ordonner les différents points, nous avons suivi la structure de la plupart des études publiées, à savoir la formulation de l'hypothèse, la méthode générale, les quantifications des mesures, l'analyse statistique et l'interprétation des résultats.

## Formulation de l'hypothèse

Une grande partie de nos connaissances en psychiatrie clinique découle d'observations rigoureuses générant des hypothèses, testées

par des études cliniques. Les investigateurs doivent s'efforcer d'énoncer précisément leur hypothèse avant de commencer, à moins que l'étude ne soit entièrement exploratoire. Dès lors, les résultats publiés des études ne correspondent pas toujours à la question formulée dans l'hypothèse, mais peuvent constituer la partie exploratoire de l'étude.

L'hypothèse, ou l'objectif, doit définir la motivation de la conduite de l'étude et influencer tant le choix du type de mesure que du schéma d'un plan pour l'analyse statistique. Généralement, l'hypothèse est énoncée dans l'introduction de la publication de l'étude.

Trop souvent, les hypothèses ne sont pas rapportées dans les études. Si les hypothèses a

Une **hypothèse** est une allégation qui peut être testée et qui concerne une relation proposée entre deux variables ou plus – allégation qui sera acceptée ou rejetée en fonction des résultats de l'étude... Si une hypothèse peut être rejetée, une autre hypothèse alternative a alors tendance à être considérée comme plus plausible. C'est-à-dire que si une hypothèse donnée est étayée, certaines interprétations sont plus probables et si l'hypothèse n'est pas étayée, d'autres interprétations sont plus probables (1).

priori ne sont pas définies, les lecteurs ne sont pas à même de discerner quelles sont les données sur lesquelles on a sélectivement insisté plus ou moins (2).

En revanche, Schooler (3) remet en question la nécessité d'une hypothèse claire. L'enchaînement subséquent dépend de la direction dans laquelle les données ont entraîné les chercheurs. D'après cet auteur, les études ne devraient pas toutes être dictées par une hypothèse. Au lieu de se concentrer sur un objectif spécifique unique qui a été fixé à l'avance, il pourrait être plus utile de dresser le meilleur tableau possible des effets des substances étudiées.

## Méthode générale

### Généralités

#### Les différents schémas

Les objectifs (hypothèses) de l'étude dictent fondamentalement le schéma. Les schémas d'étude sont intimement liés aux objectifs d'une étude et vice versa, les objectifs d'une étude doivent être liés au schéma.

Il existe différents types de schémas mais avant d'exposer les plus courants, voici la définition de certains concepts fréquemment utilisés.

Il y a d'abord la '**conception générale**' de l'étude. On peut en différencier deux types: l'étude expérimentale dans laquelle on va manipuler le facteur d'exposition et l'étude observationnelle qui laisse le champ libre à la nature, c'est-à-dire que l'on n'intervient pas dans le comportement du médecin mais on note les actions de celui-ci et l'évolution clinique du patient.

Le '**temps**' est une deuxième variable. On distingue ainsi les études statiques ou transversales, qui sont rares mais qui existent. Dans ces études, ce sont les différences à un moment précis qui sont comparées (p.ex. prévalence des posologies moyennes). Il existe aussi les études dynamiques ou longitudinales, qui sont majoritaires. L'investigateur s'intéresse à l'évolution du traitement, et des groupes de traitement peuvent être comparés à plusieurs reprises au cours de l'étude. Ces études peuvent être prospectives ou rétrospectives. Dans les premières, les données doivent encore être recueillies lors de la mise sur pied de l'étude. Généralement, une hypothèse est générée et les données sont recueillies en fonction de l'hypothèse. Dans les secondes, par contre, les données existent déjà lorsque les objectifs de l'étude sont définis. Une étude rétrospective présente l'avantage de recueillir des données du monde réel en dehors du cadre artificiel qu'une étude expérimentale peut imposer.

En outre, l'étude peut avoir ou non un haut **degré de contrôle**. Dans le cas d'un schéma hautement structuré, p.ex. les RCT, l'échantillon des patients est délimité par des critères d'inclusion et d'exclusion spécifiques. Les sujets qui satisfont aux critères constituent l'échantillon cible et sont assignés par randomisation à un groupe de traitement. Le but de la répartition aléatoire des patients sélectionnés dans différents groupes est d'obtenir des groupes identiques, sauf en ce qui concerne le médicament administré. En conséquence, le traitement administré est tenu responsable d'une différence entre les groupes. De cette manière, les

investigateurs contrôlent le mieux possible les biais. Mais lorsque le schéma est hautement structuré, le caractère généralisable est limité.

Si la randomisation est impossible parce que des groupes de grande taille ne peuvent pas être formés (batteries de tests cognitifs de grande envergure, mesures PET onéreuses), un schéma cas - témoins peut être utilisé pour résoudre le problème de l'impossibilité de randomisation, c'est-à-dire que pour chaque cas étudié on sélectionne une personne non-malade ayant un maximum de caractéristiques en commun avec le cas, p.ex. âge, sexe, éducation, habitudes, groupe socio-économique...

**L'organisation des observations** est une autre variable. On en distingue quatre types: études en simple aveugle, en double aveugle (synonyme = simple insu et double insu), études ouvertes et études croisées. Les évaluations en double aveugle contribuent à un protocole expérimental valide. On présume que les évaluations en aveugle ne sont pas biaisées par la connaissance du traitement qui est évalué. Les études de Tran et al (4) et de Conley & Mahmoud (5) sont deux exemples d'études en double aveugle. Ces études sont examinées et comparées plus loin. Dans les études ouvertes, le patient comme l'investigateur

sait quel est le médicament qui est administré au patient et le risque de biais est donc supérieur. L'étude de Ho et al (6) et l'étude QUEST (7) sont des exemples d'études ouvertes.

Les **aspects éthiques** doivent être pris en compte dans ce type de schéma. Par exemple, dans quelles circonstances est-il éthique que des individus atteints de schizophrénie, pour laquelle il existe des traitements reconnus efficaces, participent à une étude clinique? Etant donné que la thérapie psychopharmacologique est considérée comme la pierre angulaire du traitement, est-il éthiquement justifié que des sujets soient assignés à un placebo? Des patients qui présentent une réponse clinique satisfaisante au traitement, peuvent-ils être inclus et courir ainsi le risque de compromettre leur stabilité clinique? Peut-on exclure des patients gravement malades et les exposer ainsi au risque d'une aggravation clinique?

Bien sûr, les droits, la sécurité et le bien-être des sujets de l'étude sont protégés. Par exemple, les principes éthiques pour la recherche médicale impliquant des sujets humains figurent dans la "Déclaration d'Helsinki" de la *World Medical Association*. Cependant, des différences existent entre les Etats-Unis et l'Europe. Aux Etats-Unis,

#### RCT:

- hautement structurée
- la randomisation minimise la variabilité à la base
- le schéma en double aveugle contrôle les biais chez les patients, dans les évaluations de l'investigateur et dans le mode d'administration du traitement
- généralement prospective
- les biais sont minimisés autant que possible
- l'échantillon des sujets est limité aux patients qui répondent aux critères d'inclusion et consentent à leur participation à une étude clinique en double aveugle
- il se peut que le schéma ne reflète pas les besoins des patients dans les conditions du monde réel

#### Etude naturalistique:

- situations du monde réel
- échantillons représentatifs
- souvent rétrospective
- contribue à guider les médecins dans la pratique clinique
- sujette à des biais puisque le patient et le clinicien savent quel est le traitement administré (ce qui peut être exclu ultérieurement par l'analyse de la procédure de sélection)
- seuls les répondeurs font l'objet d'un rapport puisque les non-répondeurs quittent généralement l'étude précocement
- de nombreux facteurs environnementaux ne peuvent être ni identifiés ni contrôlés

l'utilisation de placebos est exigée alors qu'en Europe, les études contrôlées contre placebo sont interdites.

Les investigateurs doivent scrupuleusement déterminer si les personnes qui consentent à participer à une étude comprennent vraiment le schéma de l'étude de même que les risques et les bénéfices liés à leur participation. Les investigateurs doivent respecter le souhait des sujets de refuser de participer ou de quitter l'étude à tout moment. Généralement, on demande aux sujets de confirmer qu'ils consentent de leur plein gré à participer à une étude donnée après avoir été mis au courant de tous les aspects de l'étude qui sont pertinents pour prendre une décision sur leur participation, et ce au moyen d'un formulaire de consentement éclairé daté et signé.

Cependant, les opposants à la recherche psychiatrique argumentent que les protections actuelles comme le consentement éclairé ne sont pas suffisantes pour mettre les malades mentaux participants à l'abri de l'exploitation scientifique. Il a été démontré que la gravité des symptômes et les déficits cognitifs associés à la schizophrénie influencent négativement la capacité à donner un consentement éclairé (8).

L'application dans la pratique des résultats des RCT est souvent limitée en raison du haut degré de contrôle. Dès lors des études naturalistiques sont de plus en plus souvent effectuées. Les études naturalistiques, qui sont généralement des études rétrospectives, peuvent fournir des informations additionnelles précieuses car le schéma représente une situation du monde réel. Dès lors, les résultats des études naturalistiques peuvent contribuer à guider les médecins dans la pratique clinique. Mais comme aucun critère d'inclusion ou d'exclusion n'est spécifié et qu'aucune randomisation n'est effectuée, ce schéma peut générer davantage de biais ou de facteurs confondants. Par exemple, comme la polypharmacie ne peut être exclue, la différence entre les groupes ne peut pas être attribuée exclusivement à la différence entre les médicaments ciblés.

Tant les schémas structurés que les schémas flexibles/naturalistiques sont nécessaires pour fournir un tableau complet du mode d'utilisation des nouveaux médicaments et des médicaments existants. Si un type est confirmé par l'autre, on obtient alors un plus haut degré de certitude.

#### Validité du schéma

L'examen des différents problèmes de validité se révèle être un outil précieux pour l'interprétation des résultats d'une étude clinique et la détermination des inférences permettant de tirer des conclusions.

D'après Davidson (9), quelques facteurs fréquents qui peuvent menacer la validité globale sont:

- petit n & mesures définis antérieurement - postérieurement,
- trop de sujets quittant l'étude (*drop outs*),
- comparateur non pertinent ou inapproprié,

- courte durée de l'étude,
- schéma en aveugle déficient, mauvaise surveillance des données,
- ...

#### Problèmes spécifiques

##### Différents niveaux de comparaison

Le premier niveau est celui de la comparaison **directe** (en face-à-face) des nouveaux antipsychotiques. Il existe deux types d'études, à savoir les RCT et les études naturalistiques, que nous avons évoquées plus haut dans cet article. Dans la partie suivante, nous examinerons certains exemples d'études de comparaison directe (4-6). Il importe de comparer les résultats des RCT à ceux des études naturalistiques (10-12). Outre les études de comparaison directe, une comparaison entre l'antipsychotique et un **comparateur commun** comme un placebo ou l'halopéridol peut également être effectuée. Généralement, ces études sont des études pour

*Validité interne:* Réfère à l'efficacité du schéma, des procédures et des méthodes d'une étude pour maîtriser les variables confondantes potentielles lorsqu'on conclut que l'intervention expérimentale est responsable des différences observées entre les groupes témoin et expérimental. La validité interne est élevée dans les études RCT (critères d'inclusion rigoureux) et faible dans les études naturalistiques (vastes paramètres d'étude)

P.ex. diagnostic strict/flou, durée de la maladie, polypharmacie, etc...

*Validité externe:* la mesure dans laquelle une étude peut produire des inférences non biaisées en ce qui concerne la population cible, au-delà des sujets de l'étude (p.ex. différences entre les environnements psychosociaux contrôlés et le monde réel). La validité externe est faible dans les RCT et élevée dans les études naturalistiques.

P.ex. dans la plupart des études, les femmes sont exclues, en conséquence, les résultats ne leur sont pas nécessairement applicables (voir plus loin)

La population dont on parle dans l'étude correspond-elle à la population rencontrée dans la pratique?

*Validité de construit ou validité conceptuelle:* réfère à l'adéquation du schéma, des procédés et des mesures de résultat utilisés pour tester les hypothèses principales avec une intégrité conceptuelle. L'instrument convient-il pour mesurer ce qu'on veut savoir?

*Validité statistique:* réfère à l'utilité d'une approche pour les analyses des données en vue de détecter des différences significatives sans générer de différences factices

P.ex. correction pour les différences d'âge entre les groupes de l'étude

la FDA (*Food & Drug Administration*), ce qui signifie que des critères comparables sont imposés à toutes ces études. Bien qu'une étude de comparaison directe entre 2 médicaments procure davantage d'informations directes que la comparaison des résultats de chaque médicament versus un comparateur commun, l'approche est utile pour vérifier si certains résultats d'études de comparaison directe sont contredits s'ils proviennent d'une seule source ou tout simplement, si ces données ne sont pas disponibles. Les exemples qui illustrent ceci sont les études de Beasley et al (13) (ola vs. hal), de Marder et al (14) (ris vs. hal), de Tollefson et al (15) (ola vs. hal) et de Peuskens & Link (16) (quétiapine vs. chlorpromazine).

Un dernier type de comparaison est basé sur des expériences cliniques, **observations subjectives ou impressions** du clinicien. Par exemple, l'efficacité peut être évaluée au moyen de la documentation sur l'évolution clinique consignée par le médecin dans le dossier du patient.

### Sélection des patients

Les investigateurs doivent décrire précisément les sujets participant à une étude car les caractéristiques des patients peuvent influencer les observations. Une hétérogénéité soigneusement décrite nous permettra de faire des interprétations significatives des résultats.

Le but principal de la procédure d'échantillonnage est d'obtenir un échantillon représentatif de la population des patients en question. Des critères d'inclusion et d'exclusion bien définis doivent être soigneusement sélectionnés avant d'entreprendre toute recherche, quel que soit le schéma de l'étude. Les critères d'inclusion définissent les sujets qui présentent le diagnostic approprié, au sein de limites prédéterminées comme l'âge et d'autres facteurs (17).

Le **diagnostic** des patients enrôlés dans une étude doit être consigné et documenté (18). Ainsi, dans la plupart des études, les patients présentant des troubles apparentés à la schizophrénie sont égale-

ment autorisés à participer. Un score de **sévérité des symptômes** à la base sur la PANSS donne une description correcte et simple à la fois du groupe de patients. Les restrictions concernant le **sexe** et l'**âge** peuvent limiter l'applicabilité des résultats de l'étude provenant des patients randomisés à une population plus large. La schizophrénie est un trouble hétérogène dans le cadre duquel les individus présentent toute une série de symptômes pouvant varier au fil du temps. Par exemple, si nous ne connaissons pas l'âge de la population de l'étude, les résultats perdent de leur valeur car nous ne pouvons pas les appliquer à nos patients individuels. Par exemple, les hommes sont plus fréquemment sélectionnés pour des études cliniques car chez eux, l'âge de début de la maladie est plus précoce, la réponse aux neuroleptiques est moins bonne et l'issue est plus mauvaise (19). En outre, pour que les femmes soient incluses dans une étude clinique, il faut exclure et prévenir toute grossesse.

La **phase de la maladie** de l'échantillon peut aussi influencer les résultats des études cliniques. Les patients hospitalisés sont susceptibles de présenter une maladie grave, p.ex. un premier épisode de psychose ou une psychose rechutant de manière chronique, par rapport à des patients ambulatoires qui sont susceptibles de répondre au médicament et de fonctionner relativement bien dans la société. D'autre part, la surveillance des médicaments par le personnel de l'hôpital peut améliorer l'observance des patients hospitalisés.

Il est généralement admis que l'**exposition antérieure à des neuroleptiques** a rendu difficile l'interprétation d'un certain nombre de thèmes biologiques dans la recherche sur la schizophrénie (20). Par exemple, on peut supposer une tolérance pharmacologique après exposition antérieure prolongée. C'est pourquoi une période d'introduction sans médicament, ou période de sevrage (*wash-out*), précède l'étude. D'autre part, même après l'arrêt

complet de la prise du médicament, le délai jusqu'à la rechute du patient peut être de plusieurs mois. Enfin, la persistance d'effets indésirables après l'arrêt d'antipsychotiques explique l'incidence des SEP dans les groupes placebo par exemple.

En résumé, une étude est déterminée par toute une série de critères amenant à la sélection des patients. Parmi les critères de sélection des patients, quels sont ceux qui sont déterminants?

- le diagnostic des patients enrôlés doit être consigné et documenté,
- âge & sexe?
- comorbidité,
- polymédication et abus de substances,
- caractéristiques cliniques:
- sévérité de symptômes (PANSS),
- phase de la maladie:
- patients chroniques ou premier épisode?
- population en exacerbation aiguë ou stable?
- patients réfractaires au traitement?
- traitement antérieur?

### Taille d'échantillon

Les sujets qui satisfont aux critères d'inclusion constituent l'échantillon cible et sont issus de la population source. En conséquence, seule une petite fraction de l'échantillon cible peut véritablement être randomisée. D'habitude, ces données ne sont ni répertoriées, ni publiées.

Une fois que les critères de sélection sont définis et que les patients sont assignés par randomisation au traitement, il ne doit pas y avoir de différences statistiques intergroupes ou, si c'est le cas, des corrections statistiques doivent être apportées à ces différences.

### Taux de patients quittant l'étude

La taille de l'échantillon peut diminuer au cours de l'étude en raison des patients qui quittent l'étude. Une telle sortie d'étude concerne généralement une personne qui a décidé de changer de traitement pour de multiples raisons, y compris le manque d'efficacité, les effets indésirables et la non-observance. A leur inclusion

dans l'étude, les patients doivent généralement signer un consentement écrit. De cette manière, ils savent qu'ils participent à une étude et qu'ils courent le risque d'être traités par un placebo. Il se pourrait que cette connaissance les conduise à rompre leur engagement dans l'étude dès qu'il se produit un petit problème ou un retard dans l'apparition de la réponse. Il se pourrait en outre que les effets indésirables soient moins bien tolérés. Finalement, les taux de sorties d'étude semblent être généralement relativement élevés chez les patients schizophrènes.

Ces facteurs rendent très difficile l'interprétation de l'effet des sorties d'étude sur les résultats.

#### Adéquation du traitement

La sélection de la posologie est importante lors des comparaisons entre et dans des études de médicaments mais les critères de sélection de la dose sont mal définis. Alors que nous souhaiterions prescrire toujours la même dose d'antipsychotique, nous savons qu'une seule dose ne convient pas à tous les patients. Des patients présentant un premier épisode de schizophrénie répondent souvent à une dose d'antipsychotique plus faible que les patients qui sont malades depuis de nombreuses années. Des patients âgés ont besoin d'une dose plus faible que les jeunes patients. Des facteurs comme l'âge, la phase, la sévérité de la maladie et le traitement précédent influencent tous les mesures de la psychopathologie à la base. Un schéma d'ajustement approprié, des interactions médicamenteuses potentielles et le report des effets des médicaments précédents exerceront une influence sur la réponse (21).

Les effets du médicament peuvent mettre plus de temps à apparaître dans un domaine de symptômes que dans un autre. La mesure de l'efficacité dans tous les domaines de la schizophrénie est un problème sérieux et il est naïf de croire qu'une seule dose d'un médicament

unique agira de manière optimale dans tous les domaines. Il est également extrêmement difficile de déterminer une relation dose-réponse dans une large gamme d'effets indésirables. Le fait que les profils de liaison des nouveaux antipsychotiques exercent une influence sur plusieurs récepteurs différents rend encore plus difficile la détermination d'une relation dose-réponse (21). Un point important est de connaître la façon d'évaluer la compliance.

**La variété des modèles expérimentaux, leurs validités respectives, sont importantes. Les types d'effets, de sélection des patients, de choix de posologie, tout en tenant compte des populations étudiées, le sont tout autant.**

#### Quantification des mesures

Dans la recherche psychiatrique, il est difficile de mesurer exactement des variations dans les symptômes ou l'état clinique. C'est pourquoi des échelles soigneusement validées sont utilisées pour pouvoir se former une idée la plus précise possible sur les aspects qu'on veut mesurer. Des échelles fréquemment utilisées comme mesures du résultat sont reprises dans le tableau suivant:

Les mesures du résultat permettent de relier l'hypothèse à la conclusion ultime d'une étude. Les mesures du résultat sont utilisées pour tester l'hypothèse et pour décrire les résultats. Elles doivent refléter la réalité clinique et tester l'hypothèse énoncée.

La *validité* d'une échelle d'évaluation est le degré auquel l'échelle évalue ce qu'elle est destinée à évaluer. Par exemple un score BPRS très élevé devrait correspondre à l'évaluation d'une maladie psychotique très intense déterminée par un psychiatre expérimenté après une évaluation psychiatrique exhaustive. La *fidélité* d'une échelle est sa capacité à fournir des informations cohérentes qui sont reproductibles. La *fidélité inter-évaluateurs* est le degré de concordance entre des évaluations effectuées au même moment par différentes personnes. La *fidélité test-retest* mesure la concordance entre des évaluations effectuées à différents moments par le même évaluateur. La fidélité d'une échelle est représentée par un *coefficient*, 0 correspondant à une fidélité nulle et 1 à une fidélité parfaite. On considère généralement comme adéquats des scores supérieurs à 0,85. Le résultat doit décrire un événement pertinent

Tableau 1: Echelles fréquemment utilisées.

- Echelles donnant une impression globale sur l'état clinique:
  - CGI (*Clinical Global Impression*)
  - GAF (*Global Assessment of Functioning scale*)
- Echelles dressant un tableau plus détaillé de l'état clinique:
  - PANSS (*Positive And Negative Syndrome Scale*)
  - BPRS (*Brief Psychiatric Rating Scale*)
  - SANS (*Scale for Assessment of Negative Symptoms*)
- Echelles mesurant les effets indésirables:
  - ESRS (*Extrapyramidal Symptom Rating Scale*)
  - UKU (*Side Effect Rating Scale*)
  - BAS (*Barnes Akathisia Scale*)
  - AIMS (*Abnormal Involuntary Movement Scale*) (dyskinésie)
  - SAS (*Simpson and Angus Scale*) (Parkinsonisme)
- Echelle mesurant tous les aspects mentionnés précédemment:
  - PECC (*Psychosis Evaluation tool for Common use by Caregivers*)
- Echelles mesurant la qualité de vie:
  - SF-36 (*Health Survey*)
  - SQLS (*Quality of Life Scale*)

pour les cliniciens et pour les patients qui souffrent de la maladie. Dans de nombreuses études actuelles, des mesures du résultat évaluant le traitement à partir de la perspective du patient font défaut.

Nous avons besoin de mesures du résultat qui définissent et comptent les événements cliniques importants. Le comptage d'un événement et l'enregistrement du délai jusqu'à un événement sont des stratégies potentielles pour dépasser la variation traditionnelle en pourcentage sur une échelle d'évaluation de la base jusqu'au "point final" (*endpoint* ou dernière évaluation des résultats).

Les études cliniques évaluent l'efficacité d'une intervention alors que souvent, la pratique clinique s'efforce de prévenir des issues négatives ou de minimiser certains événements onéreux comme un séjour à l'hôpital du patient, plutôt que de viser des résultats axés sur le patient et mettant l'accent sur l'efficacité du médicament (22).

La difficulté de définir et de mesurer les résultats pose un problème. La schizophrénie est une maladie hétérogène avec différents domaines de résultats. Ces dernières années, un conflit est apparu entre le souhait de mesurer de multiples domaines de résultats (vu la complexité de la maladie et le nombre de domaines de fonctionnement différents qu'elle affecte) et l'attente qu'un ou deux schémas d'étude soient capables de répondre à toutes ces questions d'une manière interprétable et significative.

Aucune échelle d'évaluation n'est sensible au point qu'on puisse compter sur elle seule dans une étude clinique d'un médicament antipsychotique atypique. Pour illustrer ces propos, une courte présentation et comparaison de 2 échelles couramment utilisées est reprise dans l'**encadré** ci-contre.

Dans un article de Bell et al (24), la fidélité, la comparabilité et la validité prédictive de la PANSS et de la BPRS sont discutées. Bien que les deux instruments aient une haute fidélité inter-évaluateurs, les auteurs ont constaté pour la PANSS une fidélité supérieure en ce qui concerne les items de symptômes partagés et les items additionnels de la PANSS. Cette observation suggère que les descripteurs de symptômes de la PANSS sont supérieurs à ceux de la BPRS.

La PANSS a l'avantage, entre autres, de faire une distinction entre la gravité et la fréquence d'apparition d'un symptôme, des éléments qui, à la BPRS, sont combinés en un point unique. Néanmoins, la PANSS révèle toujours peu d'informations sur deux domaines importants de la schizophrénie, à savoir l'humeur et la cognition.

La cohérence interne des échelles de syndrome de la PANSS, de l'échelle générale et de l'instrument dans son ensemble s'est montrée supérieure à celle de la BPRS. Ceci suggère que les items additionnels de la PANSS et la modification des items partagés ont accru l'homogénéité des échelles de syndrome.

Une exploration de la comparabilité des items partagés a démontré que les adaptations ont suffisamment modifié la signification de ces items pour rendre la plupart d'entre eux non interchangeables.

Bien qu'elle nécessite des efforts additionnels, la PANSS produit des prédictions plus puissantes des performances de travail.

Les auteurs concluent que leurs résultats plaident en faveur de la poursuite de l'utilisation de la PANSS dans la recherche sur la schizophrénie.

La PANSS fournit un instrument d'une stabilité psychométrique et d'une puissance prédictive considérables. Pour illustrer son exactitude, nous donnons l'exemple de l'hallucination tiré du manuel d'évaluation de la PANSS (voir **figure 1**).

**Le choix des échantillons d'évaluation procure des résultats très diversifiés, allant de l'évaluation clinique globale ou détaillée, à l'observation des effets secondaires, ou encore à l'ensemble des éléments énumérés.**

## Analyse statistique

### Quelles hypothèses?

### Etudes de supériorité

Les études de supériorité sont la référence dans la comparaison de deux traitements ou

## La BPRS et la PANSS

### BPRS

La BPRS comporte 18 (originellement 16) groupements de symptômes, chacun évalué sur une échelle de sévérité de 7 points. Elle a été développée au début des années 1960 spécifiquement pour fournir une manière rapide et efficace d'évaluer les patients hospitalisés tant durant l'évaluation de médicaments que dans un contexte clinique de routine. La BPRS est l'une des plus anciennes échelles largement utilisées et elle offre l'avantage que les résultats des études utilisant la BPRS peuvent souvent être comparés à ceux d'études antérieures. Cependant, la BPRS a une limitation importante, à savoir qu'elle n'a pas été conçue pour mesurer les symptômes négatifs qui sont si importants lors de l'évaluation des antipsychotiques atypiques. Le CPMP (*Committee for proprietary medicinal products*) a recommandé l'arrêt de l'utilisation de la BPRS pour l'évaluation des patients présentant des symptômes négatifs.

### PANSS

La PANSS a été conçue comme un instrument opérationnel sensible aux médicaments qui fournit une représentation équilibrée des symptômes positifs et négatifs et évalue leur relation entre eux et à la psychopathologie globale. La PANSS développée par Kay et al (23) inclut des échelles mesurant 7 items positifs, 7 items négatifs et une échelle de 16 items de psychopathologie générale (voir liste). Au total, 18 items de la PANSS ont été adaptés de la BPRS et 12 items sont des adaptations de la *Psychopathology Rating Schedule* (PRS) de Singh & Kay. L'échelle de psychopathologie générale comporte des items comprenant 'le manque de coopération', 'l'évitement social actif' et 'le manque d'attention' qui se rapportent à une fonctionnalité individuelle mais la PANSS n'a pas de mesures directes de la capacité générale à fonctionner.



plus et utilisent l'hypothèse nulle, c'est-à-dire l'hypothèse selon laquelle il n'y a pas de différence entre les 2 groupes (ou plus) observés. L'hypothèse nulle est confrontée à l'hypothèse alternative c'est-à-dire qu'il y a une différence entre les 2 groupes (ou plus) observés. Cette différence peut alors être dans les deux sens: A est meilleur que B ou vice versa.

*bound > 0 indicated that olanzapine was statistically superior to risperidone.* "A lower bound  $\geq -4$  and  $\leq 0$  indicated there was **no difference** between the two treatments." "A lower bound  $< -4$  indicated **no conclusion** could be drawn." Dans cette étude, Tran et al (4) avaient donc comme hypothèse de départ que l'olanzapine était supérieure à la rispéridone. Les études

observée soit accidentelle ou seulement due à une coïncidence est de 5%.

### Seuil $\alpha$ ou seuil de significativité

La valeur p est comparée à un seuil de significativité (ou seuil de signification) appelé seuil  $\alpha$ , fixé arbitrairement généralement à 0,05 au début de la recherche. Si  $p < 0,05$ , on peut en conclure qu'il existe une différence statistiquement significative entre les groupes en ce qui concerne la réponse moyenne d'intérêt. Le seuil  $\alpha$  est la probabilité de conclure à tort que l'hypothèse nulle est incorrecte (= en faisant une erreur de type I c'est-à-dire la possibilité que les observations soient dues au hasard). Le seuil de signification de la valeur p est une décision arbitraire mais il est habituellement fixé à 5%.

### Tests unilatéraux et bilatéraux

Ces tests analysent les écarts par rapport à l'hypothèse.

Un test bilatéral peut analyser de tels écarts dans l'une ou l'autre direction. Il est plus valide qu'un test unilatéral et requiert une différence plus importante pour atteindre la significativité statistique. Un test unilatéral est approprié uniquement lorsque la direction de la différence entre 2 groupes est considérée comme connue avec certitude à l'avance ou lorsque des différences observées dans la direction opposée n'ont aucun intérêt ou sont impossibles. Il teste seulement les différences dans une seule direction prédéterminée et ne prend pas en considération des différences dans l'autre direction.

Les tests unilatéraux sont généralement utilisés uniquement dans un schéma de non-infériorité (voir plus haut) car ce type de schéma implique une supériorité unidirectionnelle exclusivement, qui est seule possible comme conclusion primaire.

L'étude de Tran et al (4) est un exemple d'étude de non-infériorité unilatérale et celle de Conley & Mahmoud (5), un exemple d'étude de supériorité bilatérale.

Figure 1: Sample item on "Hallucinatory behaviour" from the PANSS Rating Manual.

**Hallucinatory behavior:** Verbal report indicating perceptions which are not generated by external stimuli. These may occur in the auditory, visual, olfactory, or somatic realms. Basis for rating: verbal report and physical manifestations during the course of interview as well as report of behavior by primary care workers or family.

1. Absent – definition does not apply.
2. Minimal – questionable pathology; may be at the upper extreme of normal limits.
3. Mild – one or two clearly formed but infrequent hallucinations, or else a number of vague abnormal perceptions which do not result in distortions of thinking or behavior.
4. Moderate – hallucinations occur frequently but not continuously, and the patient's thinking and behavior are affected only to a minor extent.
5. Moderate severe – hallucinations are frequent, may involve more than one sensory modality, and tend to distort thinking and/or disrupt behavior. Patient may have a delusional interpretation of these experiences and respond to them emotionally and, on occasion, verbally as well.
6. Severe – hallucinations are present almost continuously, causing major disruption of thinking and behavior. Patient treats these as real perceptions, and functioning is impeded by frequent emotional and verbal responses to them.
7. Extreme – patient is almost totally preoccupied with hallucinations, which virtually dominate thinking and behavior. Hallucinations are provided a rigid delusional interpretation and provoke verbal and behavioral responses, including obedience to command hallucinations.

### Etudes de non-infériorité

Une étude de non-infériorité présuppose qu'un certain degré de différence est accepté comme une différence insignifiante/non pertinente. Si la différence est inférieure à ce seuil prédéterminé, le produit est considéré comme non inférieur à l'autre. Ce seuil prédéterminé est donc très important et doit être le plus minime possible tout en restant raisonnable. La conclusion primaire peut être uniquement que A est supérieur à B, ou qu'il n'est pas inférieur et donc égal à B. Seul A peut être supérieur à B et pas vice versa. Ceci est bien illustré par quelques citations de l'étude de Tran et al (4), une étude qui porte sur un test de non-infériorité de l'olanzapine versus rispéridone: "A lower

sur la non-infériorité sont davantage utilisées de nos jours.

### Valeur p

Même si les données d'une étude clinique démontrent qu'il y a une différence entre les deux groupes observés (l'hypothèse nulle est rejetée), il existe toujours l'éventualité que ce résultat observé soit dû au hasard. En d'autres termes, il se peut que, bien qu'une différence soit observée dans l'étude, il n'y ait pas de différence dans la population. Cette probabilité s'exprime par la valeur p. Plus la valeur p est petite, plus il est probable que le rejet de l'hypothèse nulle soit correct et qu'il y ait une réelle différence. Une valeur p de 0,05 indique que la probabilité que la différence

## Interprétation des résultats

### *Sélection des patients – taux de sorties d'étude*

La sélection des patients peut avoir une influence sur les résultats des études cliniques. Lorsqu'un trop petit nombre de sujets satisfait aux critères d'inclusion, la généralisabilité des résultats peut être mise en doute. Il convient de spécifier à la sélection le nombre de patients pris en considération pour se rendre compte si l'échantillon est représentatif de la population des patients en question et dans quelle mesure les résultats peuvent être généralisés. Les restrictions concernant le sexe et l'âge peuvent faire en sorte que les résultats d'étude provenant de patients randomisés ne puissent que médiocrement être extrapolés à une population plus large – par exemple l'exclusion d'un abus de substances dans les études sur la schizophrénie.

Les taux de sortie d'étude (ou taux d'abandons) doivent également être pris en compte lors de l'évaluation des résultats des études cliniques. Un échantillon qui devient trop petit peut entraîner de sérieux problèmes méthodologiques (18). Des études contrôlées contre placebo ont traditionnellement des taux d'abandons élevés puisque la plupart des patients recevant un placebo seront non-répondeurs. Les sujets sont également peu enclins à poursuivre l'utilisation d'un agent inefficace ou induisant des effets indésirables. Les notions d'abandons doivent être soigneusement précisées à l'avance et peuvent être utilisées comme variables dépendantes pour la comparaison des différents traitements (18).

### *Dose*

Lorsque 2 médicaments sont comparés dans des études cliniques, les taux de réponse de chaque médicament peuvent varier en fonction de la dose. Il est possible qu'une amélioration rapportée à une dose ne soit pas observée à d'autres doses. Inversement, une dose supérieure à la dose adéquate d'un médicament peut masquer une amélioration significative à une dose adé-

quate. C'est ainsi que nous devons garder à l'esprit que les patients ne répondent pas tous de la même manière à la même dose du médicament au cours du temps (21).

La posologie n'est pas importante 'en tant que telle' mais elle doit être prise en considération en relation avec la population de patients sélectionnée. Le protocole peut influencer la dose éventuelle.

Des études contrôlées sont toujours réalisées avec le(s) médicament(s) étudié(s) utilisé(s) en monothérapie au sein de leur classe médicamenteuse c'est-à-dire que les études avec des antipsychotiques n'associeront jamais plusieurs antipsychotiques sinon il serait impossible de tirer des conclusions. Ceci doit être pris en ligne de compte lors de l'extrapolation des résultats de l'étude à la pratique clinique au sein de laquelle la polythérapie est très fréquente (25).

Par contre, dans les études non-contrôlées, des informations fiables concernant la posologie moyenne dans l'utilisation quotidienne peuvent être récoltées.

### *Significativité statistique*

La *significativité statistique* ne doit pas être confondue avec la significativité clinique. La significativité statistique reflète l'influence du hasard sur le résultat. Si nous pouvons conclure que les résultats sont statistiquement significatifs, cela signifie que les différences observées dans ce groupe, avec ces mesures bien spécifiques, ne sont pas dues au hasard. Cela ne signifie ni plus, ni moins. La probabilité que la différence soit une coïncidence est inférieure à 5%. Le manque de significativité statistique ne veut pas dire que les 2 composés sont identiques. Cela signifie tout simplement qu'aucune différence statistique n'a été détectée pour une mesure de résultat spécifique, dans un groupe spécifique (3).

La significativité clinique reflète la *valeur clinique* du résultat. Les résultats statistiques sont tirés de groupes d'individus alors que les cliniciens traitent des individus distincts.

La mise à l'épreuve de l'hypothèse avec l'accent mis sur les valeurs p et sur la significativité peut être complétée par d'autres méthodes de rapport des résultats mais elle demeure actuellement la méthode la plus largement utilisée pour l'évaluation d'une étude clinique. Néanmoins, aucun schéma unique ne peut répondre à toutes les questions pertinentes sur un médicament et c'est l'ensemble des connaissances acquises à partir de plusieurs types de schémas qui procure les meilleures informations pour la pratique de la médecine. Chaque étude peut fournir une quantité très limitée d'informations, à savoir celles que l'étude était destinée à mesurer. Des limites peuvent apparaître sur le plan interne – des variables peuvent exercer un effet qui n'est pas reconnu, pas contrôlé ou pas mesuré – ou sur le plan externe – les extrapolations pouvant être faites à partir d'observations individuelles sont limitées par des variables comme l'âge, le sexe et la phase de la maladie. Le lecteur doit comprendre ces limitations car les décisions cliniques sont souvent basées sur des inférences tirées d'études cliniques (26). Les cliniciens sont bien servis lorsque des informations issues de plusieurs sources sont disponibles car ils peuvent ainsi mieux comprendre le mode d'action des médicaments dans le soulagement de la douleur et de la souffrance provoquées par la psychose aiguë et chronique (27).

### *Résultats probants dans des études contrôlées et naturalistiques*

Si l'hypothèse n'est pas formulée, le lecteur doit être informé que la recherche a été exploratoire. Si une telle information manque, le lecteur doit être prudent lors de l'interprétation des résultats. Les observations ne doivent certainement pas être interprétées comme définitives mais sous réserve de confirmation.

En fait, nous avons besoin tant de données dictées par une hypothèse que de données observationnelles issues d'études naturalistiques pour comprendre pleinement comment utiliser les nouveaux traitements. Si les résultats

des deux types d'études sont probants, les preuves en faveur de l'hypothèse deviennent plus solides. En d'autres termes, il est difficile de tirer des conclusions d'une seule étude, des observations comparables dans plusieurs types d'études sont nécessaires.

### *La dernière observation disponible (Last-observation-carried-forward ou LOCF)*

Lors de l'interprétation des résultats d'une étude clinique, il convient de vérifier si les abandons ont été calculés dans les effets.

Considérons par exemple une population de 100 patients susceptibles d'être traités (ITT ou *intention to treat*), mais dont seulement 80 ont été effectivement inclus dans l'étude et ont eu au moins une évaluation après la situation de base. Sur ces 80 patients, 30 ont arrêté prématurément l'étude. Il ne restait donc plus que 50 patients à la fin de l'étude. Si l'analyse est effectuée uniquement sur l'effet des patients qui ont répondu (50 patients), on peut s'attendre à un résultat plus positif que si les sorties d'étude sont également prises en compte (80 patients). On peut ainsi effectuer 2 types d'analyse: une analyse des cas observés (*observed case analysis*) qui porte sur les patients qui ont répondu et une analyse en dernière observation disponible (LOCF) au cours de laquelle la dernière mesure (parfois la seule) pour les patients ayant quitté l'étude est traitée comme une mesure de "point final" (*endpoint*). Une analyse en LOCF nous procure donc également des informations sur les patients qui n'ont pas répondu au traitement. En conséquence, une analyse en LOCF nous fournit une meilleure estimation de l'effet que le traitement exercera sur le prochain patient arbitraire. Si les deux analyses ont des résultats similaires, on peut alors être assuré que les résultats de l'analyse des cas observés ne dépendent pas des abandons. Plus l'étude est longue, plus grande est la probabilité que les observations dans les deux analyses vont différer. Dans une étude qui

dure plus d'un an, les cas observés représenteront un groupe plus finement sélectionné que dans le cas d'une étude plus courte.

### *Effets indésirables*

Lors de l'évaluation des rapports d'effets indésirables, un certain nombre d'aspects doivent être gardés à l'esprit. Avant tout, il convient de signaler avec précision si les effets indésirables ont été spontanément mentionnés par le patient ou s'ils ont été mesurés par des instruments. Une autre question pertinente est de savoir comment doivent être interprétés les effets indésirables rapportés dans le groupe placebo. Ces effets indésirables, à bien différencier de la symptomatologie clinique, sont parfois dus au traitement précédent qui a été soudainement interrompu, provoquant ainsi des symptômes de sevrage. C'est ainsi que même lors de la mesure des effets indésirables avant l'inclusion des patients dans l'étude, des SEP dus à une administration antérieure d'antipsychotiques peuvent être observés.

**Il convient de remarquer la part de l'échantillon qui a été considérée, son importance, par rapport aux abandons par exemple. De même, la dose administrée, la significativité des résultats et leur confirmation dans d'autres études doivent être prises en compte.**

## Quelques exemples de lecture critique d'études cliniques

A ce jour, seules quelques études comparant directement des antipsychotiques atypiques entre eux ont été effectuées. Dans ce paragraphe, nous commenterons deux études comparant l'olanzapine et la rispéridone (4, 5). La première de ces études a bénéficié du soutien d'*Eli Lilly and Company*, et l'autre de celui de la *Janssen Research Foundation*. Une autre étude (6) a été soutenue en partie par le NIMH (*National Institute of Mental Health*).

## Commentaires critiques

Plusieurs critiques de cette étude ont été publiées. D'après Kasper (12), le schéma de l'étude désavantage la rispéridone en ce qui concerne la posologie. La posologie choisie pour la rispéridone était basée sur les directives recommandées au moment de la conception de l'étude, à savoir jusqu'à une dose maximale de 12mg. Cette posologie ne correspond plus à la pratique clinique actuelle.

Outre ce point qui est aussi mentionné dans un article de Gheuens & Grebb (29), une analyse post-hoc de Kasper suggère qu'une série d'items diffèrent significativement entre les groupes. Le score SANS sommaire mesuré à la base était significativement supérieur dans le groupe olanzapine par rapport à celui

### **Résumé Tran et al (4) de Rogers (28)**

La première comparaison entre l'olanzapine et la rispéridone a été publiée en 1997 (4). Cette étude multicentrique prospective en double aveugle, d'une durée de 28 semaines, a inclus 339 patients atteints de schizophrénie, de trouble schizophréniforme ou de trouble schizoaffectif. Elle a en fait démarré avant la commercialisation de l'olanzapine aux Etats-Unis. Trente-cinq investigateurs dans neuf pays ont assigné par randomisation des patients soit au groupe recevant l'olanzapine (n = 172) soit au groupe recevant la rispéridone (n = 167). Les patients ont commencé le traitement par l'olanzapine à une posologie de 15mg/jour pendant les sept premiers jours de l'étude. Après cette période, les cliniciens pouvaient augmenter/diminuer la dose par paliers de 5mg tous les sept jours en restant dans une fourchette de 10 à 20mg. Les patients ont commencé l'ajustement de la dose de rispéridone à une posologie de 1mg deux fois par jour au jour 1, 2mg deux fois par jour au jour 2, après quoi la posologie a été augmentée jusqu'à 3mg deux fois par jour aux jours 3 à 7. Après la première semaine, les investigateurs étaient autorisés à augmenter/diminuer la dose par paliers de 2mg/jour tous les sept jours mais devaient demeurer dans une fourchette de 4 à 12mg/jour. La dose modale moyenne pour l'olanzapine a été de 17,2mg/jour alors que la dose modale moyenne pour la rispéridone a été de 7,2mg/jour. L'efficacité a été mesurée par la PANSS totale et ses sous-échelles, l'item de dépression de la PANSS, la BPRS de 18 items, la CGI-S et la SANS.

dans le groupe rispéridone. Ceci remet en question la différence statistiquement significative des scores SANS à l'endpoint qui a été calculée sans correction (30). Ceci ne jette pas le doute sur les conclusions concernant d'autres items.

Dans l'étude de Tran et al (31), les auteurs ont utilisé un test de non-infériorité unilatéral sans effectuer de correction statistique pour comparaisons multiples (32). Ce test a évalué l'hypothèse que l'olanzapine est supérieure à la rispéridone plutôt qu'une hypothèse nulle (c'est-à-dire qu'il n'y a pas de différences entre les médicaments). Bien que cette méthodologie soit permise pour un échantillon de plus petite taille, un test bilatéral avec des corrections aurait été plus approprié et aurait donné un poids supérieur à l'étude étant donné que c'est l'efficacité comparative des traitements qui nous intéresse plutôt que la supériorité prétendue de l'un ou de l'autre.

En outre, les auteurs rapportent qu'il n'y avait pas de différence statistiquement significative entre les 2 traitements en ce qui concerne le taux d'hospitalisation pré-étude et qu'il s'est développé une différence quant au taux d'hospitalisation post-étude. Mais ils n'indiquent pas si cette dernière différence était statistiquement significative.

Rogers (28) a conclu que tant l'olanzapine que la rispéridone ont été efficaces dans la réduction de la sévérité des symptômes

psychotiques globaux. Il est impératif de remarquer que deux valeurs p seulement sur les huit calculées comparant la variation moyenne par rapport à la base entre les groupes de traitement (c'est-à-dire la rispéridone vs l'olanzapine) ont été statistiquement significatives à savoir pour l'item de dépression de la PANSS (diminution de 1,1 pour l'olanzapine versus une diminution de 0,7 pour la rispéridone) et le score SANS sommaire (-4,3 pour l'olanzapine versus -2,9 pour la rispéridone), où il y avait déjà une différence au début.

Une incidence supérieure des symptômes extrapyramidaux (SEP), de l'hyperprolactinémie et de la dysfonction sexuelle a été notée dans le groupe de la rispéridone. Ceci peut être attribué à la dose modale moyenne élevée de 7,2mg/jour. Parmi les autres événements, seule la nausée (dans le groupe rispéridone) et le gain de poids (dans le groupe olanzapine) ont été rapportés comme statistiquement différents de l'autre groupe de traitement.

Rogers (28) donne les commentaires suivants:

Deux points à prendre en considération: la durée de cette étude (8 semaines) a été plus courte que les deux autres et la dose modale moyenne de la rispéridone dans cette étude (4,8mg/jour (ET = 1,2)) a été significativement inférieure à la dose modale moyenne de 7,2mg/jour dans l'étude de Tran et al (4).

La dose de la rispéridone dans l'étude de

Conley et Mahmoud (5) coïncide plus avec celle observée dans la pratique clinique actuelle. La dose moyenne de rispéridone à l'endpoint a été de 4,7mg/jour (ET = 1,4).

L'efficacité a été basée sur les scores à la PANSS totale et à cinq facteurs de la PANSS de même qu'à l'échelle CGI. Les auteurs ont conclu que les améliorations significatives, définies comme une réduction de > 0% à la PANSS totale, ont été observés dans les deux groupes de traitement. Une amélioration

### Abstract Ho et al (6) de Rogers (28)

Ho et al (6) ont présenté une étude ouverte dans laquelle ont été initialement enrôlés 42 patients atteints de schizophrénie (21 sous olanzapine, 21 sous rispéridone). Ils ont été suivis pendant une durée totale de 6 mois. Seuls 2/3 des sujets de l'échantillon original ont pu faire l'objet d'un suivi. Deux patients n'ont pu être réévalués lors de la conclusion de l'étude.

A la fin de l'étude, la dose moyenne était de 13,7mg/jour (DS = 4,85) pour l'olanzapine et de 6,0mg/jour (DS = 1,15) pour la rispéridone. Les effets secondaires ont été mesurés grâce aux échelles Simpson-Angus et Barnes Akathisia. Pendant le traitement aigu, les deux molécules ont été associées à une faible incidence de symptômes parkinsoniens liés au traitement, mais la rispéridone induisait plus d'akathisia. Les mesures pour le parkinsonisme ne différaient pas entre les groupes traités, même en tenant compte du taux plus élevé d'utilisation des médicaments anticholinergiques dans le groupe rispéridone.

Les mesures d'efficacité incluaient les SANS, SAPS, BPRS, GAS ainsi que les mesures de qualité de vie. Lors du suivi, les mesures globales de psychopathologie étaient significativement améliorées par rapport à l'admission dans les deux groupes.

On observait une réduction des symptômes négatifs dans les deux groupes, mais ces différences n'étaient pas significatives par rapport au niveau de départ. Le groupe rispéridone ne montrait pas de réduction significativement plus importante des symptômes psychotiques et de désorganisation. Les investigateurs concluent de ces résultats qu'il est préférable de prescrire la rispéridone aux patients présentant des délires sévères et des hallucinations. Cet effet était constaté même lorsque la dose moyenne diminuait de 5,7mg/jour à la fin de l'étude à 4,5mg/jour au 6<sup>e</sup> mois de suivi.

Jusqu'à présent, aucune critique n'a été publiée sur l'article de Ho et al (6).

### Abstract Conley & Mahmoud (5) de Rogers (28)

Conley et Mahmoud (5) ont rapporté les résultats de leur étude multicentrique, en double aveugle et randomisée, qui a évalué l'innocuité et l'efficacité de l'olanzapine et de la rispéridone chez 377 patients atteints de schizophrénie ou de trouble schizo-affectif. Les patients ont été randomisés pour recevoir quotidiennement de la rispéridone (n = 188) ou de l'olanzapine (n = 189). Le schéma posologique pour la rispéridone était de 2mg/jour aux jours 1 et 2, de 2mg/jour jusqu'à 4mg/jour aux jours 3 à 7, de 2mg/jour à 6mg/jour aux jours 8 à 14, et pour les jours 15 à 56, de 2mg/jour à 6mg/jour. Le schéma posologique pour l'olanzapine était de 10mg/jour aux jours 1 et 2, de 5mg/jour à 10mg/jour aux jours 3 à 7, de 5mg/jour à 15mg/jour aux jours 8 à 14 et de 5mg/jour à 20mg/jour aux jours 15 à 56. La dose modale moyenne de l'olanzapine était de 12,4mg/jour (ET = 4,6) alors que la dose moyenne à l'endpoint était de 13,1mg/jour (ET = 5,1).

Jusqu'à présent, aucune critique séparée n'a été publiée sur l'article de Conley et Mahmoud (5).

supérieure des symptômes positifs et des symptômes d'anxiété/dépression a été observée chez les patients recevant la rispéridone ( $p < 0,05$ ) à la semaine 8. L'olanzapine ne s'est montrée supérieure à aucune des mesures.

Cependant, les investigateurs ont rapporté qu'une incidence supérieure des événements indésirables graves a été rapportée dans le bras olanzapine par rapport au bras rispéridone (11,6% vs 8,0%). En outre, un gain de poids de 7% a été enregistré chez 27,3% des patients dans le groupe olanzapine par rapport à 11,6% des patients traités par la rispéridone.

### Résultats contradictoires ou non?

En se basant sur les données de Conley & Mahmoud (5), les examinateurs ont déterminé que les deux médicaments avaient été bien tolérés et efficaces. La rispéridone s'est révélée plus efficace pour les symptômes positifs comme pour les symptômes d'anxiété/dépression à la semaine 8. Dans l'étude de Tran et al (4), aucune différence n'a été observée entre la rispéridone et l'olanzapine pour les symptômes positifs mais les scores à l'échelle des symptômes négatifs de la SANS ont été significativement meilleurs dans le groupe olanzapine. Les auteurs ont noté que le groupe olanzapine avait un score à la base supérieur sur la SANS, c'est pourquoi la différence observée peut représenter une régression aléatoire plus importante. L'étude de Ho et al (6) étaye partiellement les observations de Conley & Mahmoud (5), à savoir que la rispéridone est plus efficace que l'olanzapine dans la réduction des symptômes négatifs après 6 mois de traitement.

Dans l'étude de Conley & Mahmoud (5), aucune différence significative n'a été constatée en ce qui concerne les symptômes extrapyramidaux, contrairement à l'étude de Tran et al (4). Comme mentionné plus haut, la dose de rispéridone utilisée par Tran et al (4) est bien supérieure à celle utilisée ces dernières années dans la pratique clinique. Puisque des doses supérieures de rispéridone peuvent induire des taux supérieurs de SEP, les résultats de l'étude de Tran et al ne sont pas surprenants. L'étude de Ho et al (6) suggère que des doses inférieures de rispéridone peuvent produire une efficacité comparable à celle de l'olanzapine, sans induire un parkinsonisme supérieur. Néanmoins, l'olanzapine a été associée à une acathisie moindre.

L'olanzapine peut exposer à certains risques sanitaires liés à sa propension à induire un gain de poids significativement supérieur dans les deux études.

Preuves probantes ou contradictoires à partir d'études ayant un comparateur commun?

Des méta-analyses peuvent fournir des mesures standardisées de l'amplitude des effets du traitement et faciliter la comparaison indirecte de nouveaux médicaments. Une méta-analyse de Leucht et al (33) résume l'efficacité et la tolérance de certains antipsychotiques récents comparés au placebo et aux antipsychotiques conventionnels. La rispéridone, l'olanzapine de même que le sertindole et la quétiapine sont tous les 4 modérément plus efficaces que le placebo. La rispéridone et l'olanzapine sont légèrement plus efficaces que l'halopéridol dans le traitement de la symptomatologie schizophrénique globale. En ce qui concerne la symptomatologie négative, tous les nouveaux antipsychotiques sont plus efficaces que le placebo. La rispéridone et l'olanzapine sont significativement supérieures à l'halopéridol. Finalement, tous les nouveaux

antipsychotiques sont associés à une utilisation moins fréquente des médicaments anti-parkinsoniens par rapport à l'halopéridol.

Les résultats variés de Tran and Conley ne sont pas contredits par les analyses des études des deux médicaments versus le placebo ou l'halopéridol.

Preuves probantes ou contradictoires à partir d'études naturalistiques?

Généralement, les résultats des études naturalistiques démontrent aussi que tant la rispéridone que l'olanzapine sont des traitements efficaces et bien tolérés pour la schizophrénie. Cependant, la rispéridone a un avantage statistiquement significatif sur l'olanzapine dans un certain nombre d'observations (10, 11, 34).

### Conclusion

La question doit être posée clairement avant qu'une réponse puisse être fournie. Les investigateurs ont la responsabilité de spécifier les hypothèses et les objectifs de l'étude. Le schéma de l'étude, la sélection des instruments d'évaluation, la population de patients, la posologie, doivent tous être orientés vers ces objectifs. Les analyses et les interprétations des données doivent être ordonnées hiérarchiquement pour répondre à l'hypothèse primaire. La responsabilité des lecteurs est de vérifier ces critères lors de la lecture des résultats.

Etant donné que les études cliniques sont vitales pour une utilisation efficace et sûre des nouveaux médicaments, il est important de comprendre les problèmes soulevés par les études, la population étudiée et les questions auxquelles l'étude répond. Il est important de garder à l'esprit qu'aucun schéma d'étude unique ne peut répondre à toutes les questions pertinentes sur les effets des médicaments antipsychotiques, même lorsqu'il s'agit d'une étude de bonne

Figure 2: Comparaison de 3 études qui comparent l'olanzapine et la rispéridone de 3 sources différentes (2 firmes pharmaceutiques et académique).

	Tran et al (4)	Conley & Mahmoud (5)	Ho et al (6)
Objectif	Montrer la supériorité OLA sur RIS pour les caractéristiques d'innocuité et/ou d'efficacité	Comparaison entre RIS et OLA pour les taux de SEP et l'efficacité	Comparaison entre RIS et OLA pour SEP et efficacité
Mesures du résultat	PANSS, BPRS, CGI, SANS, Simpson-Angus, Barnes Akathisia, AIMS, QoL	PANSS, ESRS, poids corporel, indice de masse corporelle	BPRS, SANS, SAPS, GAS, QoL (Psych-Base), Simpson-Angus, Barnes Akathisia
Méthode statistique	Test t unilatéral Etude de non-infériorité	Test t bilatéral Etude de supériorité	Test t bilatéral Étude de supériorité
Seuil $\alpha$	0,05	0,05	0,05
Evaluation	Aveugle	Aveugle	Ouverte
Recrutement	Randomisé	Randomisé	Alternativement naturalistique
Posologie	RIS 4-12mg/jour (7,2 $\pm$ 2,7) OLA 10-20mg/jour (17,2 $\pm$ 3,6)	RIS 2-6mg/jour (4,8mg) OLA 10-20mg/jour (12,4mg)	A la sortie d'hôpital: RIS: 5,7mg/jour OLA: 14,4mg/jour Au suivi: RIS: 4,5mg/jour OLA: 13,8mg/jour
Diagnostic	81,7% de patients schizophréniques DSM	86,2% de patients schizophréniques DSM	100% de patients schizophréniques DSM
Age	âge moyen: 36,21	Age moyen: 40,0 le groupe olanzapine (38,9) était légèrement plus jeune que le groupe de la rispéridone (41,0)	Groupe ola: âge moyen: 33,5 Groupe ris: âge moyen: 29,6
Sexe	64,9% d'hommes	72,7% d'hommes	76,2% d'hommes
Patients	hospitalisés (100%)	hospitalisés (21%) et ambulatoires (79%)	hospitalisés (100%)
Gravité des symptômes	min. de 42 à la BPRS	Score PANSS totale à la base $\geq$ 60 et $\leq$ 120	BPRS totale RIS: 46,3 OLA: 43,9
Abandons	42,4% dans le groupe olanzapine 52,7% dans le groupe rispéridone	23% dans le groupe olanzapine 28% dans le groupe rispéridone	
Traitement précédent	Non rapporté	$\pm$ 50% antipsychotiques 15% AP conventionnels	- naïfs aux neuroleptiques - arrêt antérieur du traitement - consentement au retrait
Conclusions	innocuité: RIS = OLA efficacité: OLA > RIS sympt. nég. SANS/sympt. dépression PANSS) effets indésirables: RIS > OLA gain de poids: OLA > RIS valeurs de laboratoire: prolactine: OLA meilleure que RIS Concentration transaminases: OLA plus élevée que RIS mais pas à l'endpoint	innocuité: RIS = OLA efficacité: RIS > OLA sur sympt. pos. et anx/dépr (semaine 8) SEP: RIS = OLA Gain de poids: OLA > RIS Valeurs lab.: profils transaminases hépatiques et lipides: OLA plus grave que RIS prolactine: RIS plus grave que OLA	efficacité: RIS = OLA sur réduction sympt. nég., pos. et sympt. désorg. durant le traitement aigu, RIS > OLA sur réduction sympt. pos. à 6 mois acathisie: OLA meilleure que RIS

qualité et correcte dans laquelle toutes les variables confondantes ont été exclues ou contrôlées. Ces études conduisent souvent à un nombre limité de patients issus de la population source. C'est une combinaison de plusieurs études qui informe le mieux les médecins. Bien que leur concept soit moins contrôlé, les études naturalistiques peuvent toutefois être utiles. Les différentes approches doivent concorder.

### **Directives: les aspects les plus importants lors de l'évaluation ou de l'interprétation d'une étude:**

#### 1. Sélection des patients:

- Le groupe des patients sélectionnés est-il similaire aux patients que vous rencontrez dans votre pratique (diagnostic, âge, sexe, statut de l'hospitalisation)?
- Si non, quelles sont les différences?

#### 2. Posologie:

- La posologie utilisée dans l'étude est-elle pertinente pour votre population de patients?
- Si oui, les résultats de l'étude concordent-ils avec votre expérience?

#### 3. Résultats:

- Significativité statistique
- Significativité clinique

#### 4. Généralisation

- Quelles sont les raisons d'une faible généralisation estimée?
- L'utilisation de polythérapies dans la pratique clinique peut conduire à des expériences qui ne concordent pas avec les résultats de l'étude, puisque la plupart des études sont effectuées en monothérapie
- Les extrapolations des observations individuelles sont limitées par les variables telles que l'âge, le sexe, la phase de la maladie, l'abus de substances...

#### 5. Recherche de confirmation par d'autres études contrôlées et naturalistiques

#### 6. Adéquation de la méthodologie

## Références

1. Lang TA, Secic M. How to report statistics in Medicine: Annotated guidelines for authors, editors and reviewers. Philadelphia: American College of Physicians; 1997.
2. Lieberman JA. Hypothesis and hypothesis testing in clinical trial. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):5-10.
3. Schooler NR. The statistical comparison of clinical trials. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):35-9.
4. Tran PV, Hamilton SH, Kuntz AJ, et al. Double-blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders. *Journal of Clinical Psychopharmacology* 1997;17(5):407-18.
5. Conley RR, Mahmoud R. A randomized double-blind study of risperidone and olanzapine in the treatment of schizophrenia or schizoaffective disorder. *American Journal of Psychiatry* 2001;158(5):765-74.
6. Ho BC, Miller D, Nopoulos P, Andreasen NC. A comparative effectiveness study of risperidone and olanzapine in the treatment of schizophrenia. *Journal of Clinical Psychiatry* 1999;60(10):658-63.
7. Mullen J, Reinstein M, Bari M, et al. Quetiapine and risperidone in outpatients with psychotic disorders: results of the QUEST trial. Presented at the International Congress on Schizophrenia Research Biennial Meeting. 1999.
8. Roberts, L. W. Ethics of psychiatric research: conceptual issues and empirical findings. *Comprehensive Psychiatry* 1998;39:99-110.
9. Davidson. *Communication Personelle*. 1999.
10. Procyshyn RM, Zerjav S. Drug utilization patterns and outcomes associated with in-hospital treatment with risperidone of olanzapine. *Clinical Therapeutics* 1998;20(6):1203-17.
11. Snatser M, Welch R. A retrospective, naturalistic review comparing clinical outcomes of in-hospital treatment with risperidone and olanzapine. *Clin Drug Invest* 2000;20(3):159-64.
12. Kasper S. Comments on "double blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders" by Tran and Associates. *Journal of Clinical Psychopharmacology* 1998;18(4):353-4.
13. Beasley CM, Tollefson G, Tran P, Satterlee W, Sanger T, Hamilton S. Olanzapine versus placebo and haloperidol: acute phase results of the North American double-blind olanzapine trial. *Neuropsychopharmacology*. 1996a;14(2):111-23.
14. Marder SR, Davis JM, Chouinard G. The Effects of Risperidone on the Five Dimensions of Schizophrenia Derived by Factor Analysis: Combined Results of the North American Trials. *Journal of Clinical Psychiatry* 1997;58:538-46.
15. Tollefson GD, Beasley CM Jr, Tran PV, et al. Olanzapine versus haloperidol in the treatment of schizophrenia and schizoaffective and schizophreniform disorders: results of an international collaborative trial. *American Journal of Psychiatry* 1997;154:457-65.
16. Peuskens J, Link CGG. A comparison of quetiapine and chlorpromazine in the treatment of schizophrenia. *Acta Psychiatrica Scandinavica* 1997;96:265-73.
17. Licht RW, Gouliaev G, Vestergaard P, et al. Generalizability of results from randomized drug trials: a trial on antimanic treatment. *British Journal of Psychiatry* 1997;170:264-7.
18. Angst J, Bech P, Bobon D, et al. Report on the Third Consensus Conference on the Methodology of Clinical Trials With Antipsychotic Drugs. *Pharmacopsychiatry* 1991;24:149-52.
19. Lindamer LA, Lohr JB, Harris MJ, et al. Gender-related clinical differences in older patients with schizophrenia. *Journal of Clinical Psychiatry* 1999;60:61-7.
20. Kirch DG, Keith SJ, Matthews S. M. Research on first episode psychosis: report on a National Institute of Mental Health Workshop. *Schizophrenia Bulletin*. 1992; 18:179-84.
21. Kane JM. Dose selection and comparator drugs in schizophrenia research. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):29-34.
22. Conley RR. Evaluating Clinical trial data: outcome measures. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):23-8.
23. Kay SR, Opler LA, Lindenmayer JP. The positive and negative syndrome scale (PANSS): rationale and standardisation. *British Journal of Psychiatry* 1989;155(7):59.
24. Bell M, Milstein R, Beam-Goulet J, Lysaker P, Cicchetti D. The positive and negative syndrome scale and the brief psychiatric rating scale. *Journal of Nervous and Mental Disease*. 1992;180(11):723-8.
25. De Hert M. Psychose schizophrénique et abus de substances. Dom G, editor. *Troubles psychiques et accoutumance. Comment les traiter ensemble?* EPO 2000. p. 221-44.
26. Docherty JP. Interpretations and conclusions in the clinical trial. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):40-5.
27. Casey DE. Clinical trial design issues in schizophrenic research. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):17-22.
28. Rogers D. Comparing atypical antipsychotics. *Psychiatric Times* 2002;19(1).
29. Gheuens JA, Grebb J. Comments on article by Tran and associates, 'Double-blind comparison of olanzapine versus risperidone in treatment of schizophrenia and other psychotic disorders'. *Journal of Clinical Psychopharmacology* 1998;18:176-9.
30. Keith SJ. Evaluating Characteristics of patient selection and dropout rates. *Journal of Clinical Psychiatry* 2001;62(Suppl.9):11-6.
31. Tran PV, Dellva MA, Tollefson GD, Wentley AL, Beasley CM. Oral olanzapine versus oral haloperidol in the maintenance treatment of schizophrenia and related psychosis. *British Journal of Psychiatry* 1998;172:499-505.
32. Schooler N. Comments on Article by Tran and Colleagues, "double blind comparison of olanzapine versus risperidone in treatment of schizophrenia and other psychotic disorders". *Journal of Clinical Psychopharmacology* 1998;18(2):174-5.
33. Leucht S, Pitschel-Walz G, Abraham D, Kissling W. Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo. A meta-analysis of randomized controlled trials. *Schizophrenia Research* 1999;35(1):51-68.
34. Kasper S, Rosillon D, Duchesne I. Risperidone olanzapine drug outcomes studies in schizophrenia (RODOS): efficacy and tolerability results of an international naturalistic study. *International Clinical Psychopharmacology* 2001;16:179-87.

The work of the Belgian Discussion Board on Efficacy of Antipsychotic Treatment is supported by an unrestricted educational grant from Janssen-Cilag & Organon.