



HAL
open science

**Fidélité inter-juges d'un instrument de mesure des
symptômes prémorbides et prodromiques de la
schizophrénie. A propos d'une étude rétrospective
réalisée à Nancy sur 50 dossiers**

Félix Arrondo

► **To cite this version:**

Félix Arrondo. Fidélité inter-juges d'un instrument de mesure des symptômes prémorbides et prodromiques de la schizophrénie. A propos d'une étude rétrospective réalisée à Nancy sur 50 dossiers. Sciences du Vivant [q-bio]. 2012. hal-01733147

HAL Id: hal-01733147

<https://hal.univ-lorraine.fr/hal-01733147>

Submitted on 14 Mar 2018

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-thesesexercice-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

THESE

pour obtenir le grade de

DOCTEUR EN MEDECINE

Présentée et soutenue publiquement
dans le cadre du troisième cycle de Médecine Spécialisée

par

Félix ARRONDO

le 14 juin 2012

**FIDELITE INTER-JUGES D'UN INSTRUMENT DE MESURE DES SYMPTOMES
PREMORBIDES ET PRODROMIQUES DE LA SCHIZOPHRENIE**

A propos d'une étude rétrospective réalisée à Nancy sur 50 dossiers

Examineurs de la thèse :

M. B. KABUTH	Professeur	Président
Mme C. VIDAILHET	Professeur	Juge
M. D. SIBERTIN-BLANC	Professeur	Juge
Mme C. OBACZ	Docteur	Juge et Directrice

UNIVERSITÉ DE LORRAINE
FACULTÉ DE MÉDECINE DE NANCY

Administrateur Provisoire de l'Université de Lorraine : Professeur Jean-Pierre FINANCE

Doyen de la Faculté de Médecine : Professeur Henry COUDANE

Vice Doyen « Pédagogie » : Professeur Karine ANGIOI
Vice Doyen *Mission* « *sillon lorrain* » : Professeur Annick BARBAUD
Vice Doyen *Mission* « *Campus* » : Professeur Marie-Christine BÉNÉ
Vice Doyen *Mission* « *Finances* » : Professeur Marc BRAUN
Vice Doyen *Mission* « *Recherche* » : Professeur Jean-Louis GUÉANT

Asseseurs :

- 1 ^{er} Cycle :	Professeur Bruno CHENUÉL
- « Première année commune aux études de santé (PACES) et universitarisation études para-médicales »	M. Christophe NÉMOS
- 2 ^{ème} Cycle :	Professeur Marc DEBOUVERIE
- 3 ^{ème} Cycle :	
« <i>DES Spécialités Médicales, Chirurgicales et Biologiques</i> »	Professeur Jean-Pierre BRONOWICKI
« <i>DES Spécialité Médecine Générale</i> »	Professeur Francis RAPHAËL
- Filières professionnalisées :	M. Walter BLONDEL
- Formation Continue :	Professeur Hervé VESPIGNANI
- Commission de Prospective :	Professeur Pierre-Edouard BOLLAERT
- Recherche :	Professeur Didier MAINARD
- Développement Professionnel Continu :	Professeur Jean-Dominique DE KORWIN
Asseseurs Relations Internationales	Professeur Jacques HUBERT

DOYENS HONORAIRES

Professeur Adrien DUPREZ – Professeur Jean-Bernard DUREUX
Professeur Jacques ROLAND – Professeur Patrick NETTER

=====

PROFESSEURS HONORAIRES

Pierre ALEXANDRE – Jean-Marie ANDRE - Daniel ANTHOINE - Alain BERTRAND - Pierre BEY – Patrick BOISSEL
Jacques BORRELLY - Michel BOULANGE - Jean-Claude BURDIN - Claude BURLET - Daniel BURNEL
Claude CHARDOT - Jean-Pierre CRANCE - Gérard DEBRY - Jean-Pierre DELAGOUTTE - Emile de LAVERGNE
Jean-Pierre DESCHAMPS - Jean DUHEILLE - Adrien DUPREZ - Jean-Bernard DUREUX - Gérard FIEVE Jean
FLOQUET - Robert FRISCH - Alain GAUCHER - Pierre GAUCHER - Hubert GERARD
Jean-Marie GILGENKRANTZ - Simone GILGENKRANTZ - Oliéro GUERCI - Pierre HARTEMANN - Claude HURIET
Christian JANOT – Michèle KESSLER - Jacques LACOSTE - Henri LAMBERT - Pierre LANDES
Marie-Claire LAXENAIRE - Michel LAXENAIRE - Jacques LECLERE - Pierre LEDERLIN - Bernard LEGRAS
Michel MANCIAUX - Jean-Pierre MALLIÉ – Philippe MANGIN - Pierre MATHIEU - Denise MONERET-VAUTRIN
Pierre MONIN - Pierre NABET - Jean-Pierre NICOLAS - Pierre PAYSANT - Francis PENIN - Gilbert PERCEBOIS
Claude PERRIN - Guy PETIET - Luc PICARD - Michel PIERSON - Jean-Marie POLU – Jacques POUREL
Jean PREVOT - Antoine RASPILLER - Michel RENARD - Jacques ROLAND - René-Jean ROYER - Daniel SCHMITT
Michel SCHWEITZER – Claude SIMON - Danièle SOMMELET – Jean-François STOLTZ
Michel STRICKER - Gilbert THIBAUT - Augusta TREHEUX - Hubert UFFHOLTZ - Gérard VAILLANT - Paul VERT
Colette VIDAILHET - Michel VIDAILHET - Michel WAYOFF - Michel WEBER

=====

**PROFESSEURS DES UNIVERSITÉS
PRATICIENS HOSPITALIERS**

(Disciplines du Conseil National des Universités)

42^{ème} Section : MORPHOLOGIE ET MORPHOGENÈSE

1^{ère} sous-section : (*Anatomie*)

Professeur Gilles GROSDIDIER

Professeur Pierre LASCOMBES – Professeur Marc BRAUN

2^{ème} sous-section : *(Cytologie et histologie)*

Professeur Bernard FOLIGUET

3^{ème} sous-section : *(Anatomie et cytologie pathologiques)*

Professeur François PLENAT – Professeur Jean-Michel VIGNAUD

43^{ème} Section : BIOPHYSIQUE ET IMAGERIE MÉDICALE

1^{ère} sous-section : *(Biophysique et médecine nucléaire)*

Professeur Gilles KARCHER – Professeur Pierre-Yves MARIE – Professeur Pierre OLIVIER

2^{ème} sous-section : *(Radiologie et imagerie médicale)*

Professeur Denis REGENT – Professeur Michel CLAUDON – Professeur Valérie CROISÉ-LAURENT

Professeur Serge BRACARD – Professeur Alain BLUM – Professeur Jacques FELBLINGER

Professeur René ANXIONNAT

44^{ème} Section : BIOCHIMIE, BIOLOGIE CELLULAIRE ET MOLÉCULAIRE, PHYSIOLOGIE ET NUTRITION

1^{ère} sous-section : *(Biochimie et biologie moléculaire)*

Professeur Jean-Louis GUÉANT – Professeur Jean-Luc OLIVIER – Professeur Bernard NAMOUR

2^{ème} sous-section : *(Physiologie)*

Professeur François MARCHAL – Professeur Bruno CHENUÉL – Professeur Christian BEYAERT

3^{ème} sous-section : *(Biologie Cellulaire)*

Professeur Ali DALLLOUL

4^{ème} sous-section : *(Nutrition)*

Professeur Olivier ZIEGLER – Professeur Didier QUILLIOT - Professeur Rosa-Maria RODRIGUEZ-GUEANT

45^{ème} Section : MICROBIOLOGIE, MALADIES TRANSMISSIBLES ET HYGIÈNE

1^{ère} sous-section : *(Bactériologie – virologie ; hygiène hospitalière)*

Professeur Alain LE FAOU - Professeur Alain LOZNIEWSKI

3^{ème} sous-section : *(Maladies infectieuses ; maladies tropicales)*

Professeur Thierry MAY – Professeur Christian RABAUD

46^{ème} Section : SANTÉ PUBLIQUE, ENVIRONNEMENT ET SOCIÉTÉ

1^{ère} sous-section : *(Épidémiologie, économie de la santé et prévention)*

Professeur Philippe HARTEMANN – Professeur Serge BRIANÇON - Professeur Francis GUILLEMIN

Professeur Denis ZMIROU-NAVIER – Professeur François ALLA

2^{ème} sous-section : *(Médecine et santé au travail)*

Professeur Christophe PARIS

3^{ème} sous-section : *(Médecine légale et droit de la santé)*

Professeur Henry COUDANE

4^{ème} sous-section : *(Biostatistiques, informatique médicale et technologies de communication)*

Professeur François KOHLER – Professeur Éliane ALBUISSON

47^{ème} Section : CANCÉROLOGIE, GÉNÉTIQUE, HÉMATOLOGIE, IMMUNOLOGIE

1^{ère} sous-section : *(Hématologie ; transfusion)*

Professeur Pierre BORDIGONI - Professeur Jean-François STOLTZ – Professeur Pierre FEUGIER

2^{ème} sous-section : *(Cancérologie ; radiothérapie)*

Professeur François GUILLEMIN – Professeur Thierry CONROY

Professeur Didier PEIFFERT – Professeur Frédéric MARCHAL

3^{ème} sous-section : *(Immunologie)*

Professeur Gilbert FAURE – Professeur Marie-Christine BENE

4^{ème} sous-section : *(Génétique)*

Professeur Philippe JONVEAUX – Professeur Bruno LEHEUP

**48^{ème} Section : ANESTHÉSIOLOGIE, RÉANIMATION, MÉDECINE D'URGENCE,
PHARMACOLOGIE ET THÉRAPEUTIQUE**

1^{ère} sous-section : *(Anesthésiologie et réanimation chirurgicale ; médecine d'urgence)*

Professeur Claude MEISTELMAN – Professeur Hervé BOUAZIZ

Professeur Paul-Michel MERTES – Professeur Gérard AUDIBERT – Professeur Thomas FUCHS-BUDER

2^{ème} sous-section : *(Réanimation médicale ; médecine d'urgence)*

Professeur Alain GERARD - Professeur Pierre-Édouard BOLLAERT

Professeur Bruno LÉVY – Professeur Sébastien GIBOT

3^{ème} sous-section : *(Pharmacologie fondamentale ; pharmacologie clinique ; addictologie)*

Professeur Patrick NETTER – Professeur Pierre GILLET

4^{ème} sous-section : *(Thérapeutique ; médecine d'urgence ; addictologie)*

Professeur François PAILLE – Professeur Gérard GAY – Professeur Faiez ZANNAD - Professeur Patrick ROSSIGNOL

49^{ème} Section : PATHOLOGIE NERVEUSE ET MUSCULAIRE, PATHOLOGIE MENTALE, HANDICAP et RÉÉDUCATION

1^{ère} sous-section : (Neurologie)

Professeur Gérard BARROCHE – Professeur Hervé VESPIGNANI
Professeur Xavier DUCROCQ – Professeur Marc DEBOUVERIE – Professeur Luc TAILLANDIER

2^{ème} sous-section : (Neurochirurgie)

Professeur Jean-Claude MARCHAL – Professeur Jean AUQUE – Professeur Olivier KLEIN
Professeur Thierry CIVIT

3^{ème} sous-section : (Psychiatrie d'adultes ; addictologie)

Professeur Jean-Pierre KAHN – Professeur Raymund SCHWAN

4^{ème} sous-section : (Pédopsychiatrie ; addictologie)

Professeur Daniel SIBERTIN-BLANC – Professeur Bernard KABUTH

5^{ème} sous-section : (Médecine physique et de réadaptation)

Professeur Jean PAYSANT

50^{ème} Section : PATHOLOGIE OSTÉO-ARTICULAIRE, DERMATOLOGIE et CHIRURGIE PLASTIQUE

1^{ère} sous-section : (Rhumatologie)

Professeur Isabelle CHARY-VALCKENAERE – Professeur Damien LOEUILLE

2^{ème} sous-section : (Chirurgie orthopédique et traumatologique)

Professeur Daniel MOLE - Professeur Didier MAINARD

Professeur François SIRVEAUX – Professeur Laurent GALOIS

3^{ème} sous-section : (Dermato-vénéréologie)

Professeur Jean-Luc SCHMUTZ – Professeur Annick BARBAUD

4^{ème} sous-section : (Chirurgie plastique, reconstructrice et esthétique ; brûlologie)

Professeur François DAP – Professeur Gilles DAUTEL

51^{ème} Section : PATHOLOGIE CARDIORESPIRATOIRE et VASCULAIRE

1^{ère} sous-section : (Pneumologie ; addictologie)

Professeur Yves MARTINET – Professeur Jean-François CHABOT – Professeur Ari CHAOUAT

2^{ème} sous-section : (Cardiologie)

Professeur Etienne ALIOT – Professeur Yves JUILLIERE – Professeur Nicolas SADOUL

Professeur Christian de CHILLOU

3^{ème} sous-section : (Chirurgie thoracique et cardiovasculaire)

Professeur Jean-Pierre VILLEMOT

4^{ème} sous-section : (Chirurgie vasculaire ; médecine vasculaire)

Professeur Denis WAHL – Professeur Sergueï MALIKOV

52^{ème} Section : MALADIES DES APPAREILS DIGESTIF et URINAIRE

1^{ère} sous-section : (Gastroentérologie ; hépatologie ; addictologie)

Professeur Marc-André BIGARD - Professeur Jean-Pierre BRONOWICKI – Professeur Laurent PEYRIN-BIROULET

3^{ème} sous-section : (Néphrologie)

Professeur Dominique HESTIN – Professeur Luc FRIMAT

4^{ème} sous-section : (Urologie)

Professeur Jacques HUBERT – Professeur Pascal ESCHWEGE

53^{ème} Section : MÉDECINE INTERNE, GÉRIATRIE et CHIRURGIE GÉNÉRALE

1^{ère} sous-section : (Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale ; addictologie)

Professeur Jean-Dominique DE KORWIN – Professeur Pierre KAMINSKY

Professeur Athanase BENETOS - Professeur Gisèle KANNY – Professeur Christine PERRET-GUILLAUME

2^{ème} sous-section : (Chirurgie générale)

Professeur Laurent BRESLER - Professeur Laurent BRUNAUD – Professeur Ahmet AYAV

54^{ème} Section : DÉVELOPPEMENT ET PATHOLOGIE DE L'ENFANT, GYNÉCOLOGIE-OBSTÉTRIQUE, ENDOCRINOLOGIE ET REPRODUCTION

1^{ère} sous-section : (Pédiatrie)

Professeur Jean-Michel HASCOET - Professeur Pascal CHASTAGNER

Professeur François FEILLET - Professeur Cyril SCHWEITZER – Professeur Emmanuel RAFFO

2^{ème} sous-section : (Chirurgie infantile)

Professeur Michel SCHMITT – Professeur Pierre JOURNEAU – Professeur Jean-Louis LEMELLE

3^{ème} sous-section : (Gynécologie-obstétrique ; gynécologie médicale)

Professeur Jean-Louis BOUTROY - Professeur Philippe JUDLIN – Professeur Patricia BARBARINO

4^{ème} sous-section : (Endocrinologie, diabète et maladies métaboliques ; gynécologie médicale)

Professeur Georges WERYHA – Professeur Marc KLEIN – Professeur Bruno GUERCI

55^{ème} Section : PATHOLOGIE DE LA TÊTE ET DU COU

1^{ère} sous-section : (*Oto-rhino-laryngologie*)

Professeur Roger JANKOWSKI – Professeur Cécile PARIETTI-WINKLER

2^{ème} sous-section : (*Ophtalmologie*)

Professeur Jean-Luc GEORGE – Professeur Jean-Paul BERROD – Professeur Karine ANGIOI-DUPREZ

3^{ème} sous-section : (*Chirurgie maxillo-faciale et stomatologie*)

Professeur Jean-François CHASSAGNE – Professeur Etienne SIMON

=====

PROFESSEURS DES UNIVERSITÉS

64^{ème} Section : BIOCHIMIE ET BIOLOGIE MOLÉCULAIRE

Professeur Sandrine BOSCHI-MULLER

PROFESSEURS DES UNIVERSITÉS DE MÉDECINE GÉNÉRALE

Professeur Jean-Marc BOIVIN

=====

MAÎTRES DE CONFÉRENCES DES UNIVERSITÉS - PRATICIENS HOSPITALIERS

42^{ème} Section : MORPHOLOGIE ET MORPHOGENÈSE

1^{ère} sous-section : (*Anatomie*)

Docteur Bruno GRIGNON – Docteur Thierry HAUMONT – Docteur Manuela PEREZ

2^{ème} sous-section : (*Cytologie et histologie*)

Docteur Edouard BARRAT - Docteur Françoise TOUATI – Docteur Chantal KOHLER

3^{ème} sous-section : (*Anatomie et cytologie pathologiques*)

Docteur Aude BRESSENOT

43^{ème} Section : BIOPHYSIQUE ET IMAGERIE MÉDICALE

1^{ère} sous-section : (*Biophysique et médecine nucléaire*)

Docteur Marie-Hélène LAURENS – Docteur Jean-Claude MAYER

Docteur Pierre THOUVENOT – Docteur Jean-Marie ESCANYE

2^{ème} sous-section : (*Radiologie et imagerie médicale*)

Docteur Damien MANDRY

44^{ème} Section : BIOCHIMIE, BIOLOGIE CELLULAIRE ET MOLÉCULAIRE, PHYSIOLOGIE ET NUTRITION

1^{ère} sous-section : (*Biochimie et biologie moléculaire*)

Docteur Jean STRACZEK – Docteur Sophie FREMONT

Docteur Isabelle GASTIN – Docteur Marc MERTEN – Docteur Catherine MALAPLATE-ARMAND

Docteur Shyue-Fang BATTAGLIA

3^{ème} sous-section : (*Biologie Cellulaire*)

Docteur Véronique DECOT-MAILLERET

45^{ème} Section : MICROBIOLOGIE, MALADIES TRANSMISSIBLES ET HYGIÈNE

1^{ère} sous-section : (*Bactériologie – Virologie ; hygiène hospitalière*)

Docteur Francine MORY – Docteur Véronique VENARD – Docteur Hélène JEULIN

2^{ème} sous-section : (*Parasitologie et mycologie*)

Docteur Nelly CONTET-AUDONNEAU – Madame Marie MACHOUART

46^{ème} Section : SANTÉ PUBLIQUE, ENVIRONNEMENT ET SOCIÉTÉ

1^{ère} sous-section : (*Epidémiologie, économie de la santé et prévention*)

Docteur Alexis HAUTEMANIÈRE – Docteur Frédérique CLAUDOT – Docteur Cédric BAUMANN

2^{ème} sous-section (*Médecine et Santé au Travail*)

Docteur Isabelle THAON

3^{ème} sous-section (*Médecine légale et droit de la santé*)

Docteur Laurent MARTRILLE

4^{ème} sous-section : (*Biostatistiques, informatique médicale et technologies de communication*)

Docteur Nicolas JAY

47^{ème} Section : CANCÉROLOGIE, GÉNÉTIQUE, HÉMATOLOGIE, IMMUNOLOGIE
2^{ème} sous-section : (Cancérologie ; radiothérapie : cancérologie (type mixte : biologique)

Docteur Lina BOLOTINE

3^{ème} sous-section : (Immunologie)

Docteur Marcelo DE CARVALHO BITTENCOURT

4^{ème} sous-section : (Génétique)

Docteur Christophe PHILIPPE – Docteur Céline BONNET

**48^{ème} Section : ANESTHÉSIOLOGIE, RÉANIMATION, MÉDECINE D'URGENCE,
PHARMACOLOGIE ET THÉRAPEUTIQUE**

3^{ème} sous-section : (Pharmacologie fondamentale ; pharmacologie clinique)

Docteur Françoise LAPICQUE – Docteur Marie-José ROYER-MORROT – Docteur Nicolas GAMBIER

50^{ème} Section : RHUMATOLOGIE

1^{ère} sous-section : (Rhumatologie)

Docteur Anne-Christine RAT

3^{ème} sous-section : (Dermato-vénéréologie)

Docteur Anne-Claire BURSZTEJN

53^{ème} Section : MÉDECINE INTERNE, GÉRIATRIE et CHIRURGIE GÉNÉRALE

1^{ère} sous-section : (Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale ; addictologie)

Docteur Laure JOLY

**54^{ème} Section : DÉVELOPPEMENT ET PATHOLOGIE DE L'ENFANT, GYNÉCOLOGIE-OBSTÉTRIQUE,
ENDOCRINOLOGIE ET REPRODUCTION**

3^{ème} sous-section :

Docteur Olivier MOREL

5^{ème} sous-section : (Biologie et médecine du développement et de la reproduction ; gynécologie médicale)

Docteur Jean-Louis CORDONNIER

=====

MAÎTRE DE CONFÉRENCE DES UNIVERSITÉS DE MÉDECINE GÉNÉRALE

Docteur Elisabeth STEYER

=====

MAÎTRES DE CONFÉRENCES

5^{ème} section : SCIENCE ÉCONOMIE GÉNÉRALE

Monsieur Vincent LHUILLIER

40^{ème} section : SCIENCES DU MÉDICAMENT

Monsieur Jean-François COLLIN

60^{ème} section : MÉCANIQUE, GÉNIE MÉCANIQUE ET GÉNIE CIVILE

Monsieur Alain DURAND

61^{ème} section : GÉNIE INFORMATIQUE, AUTOMATIQUE ET TRAITEMENT DU SIGNAL

Monsieur Jean REBSTOCK – Monsieur Walter BLONDEL

64^{ème} section : BIOCHIMIE ET BIOLOGIE MOLÉCULAIRE

Mademoiselle Marie-Claire LANHERS – Monsieur Pascal REBOUL – Mr Nick RAMALANJAONA

65^{ème} section : BIOLOGIE CELLULAIRE

Mademoiselle Françoise DREYFUSS – Monsieur Jean-Louis GELLY
Madame Ketsia HESS – Monsieur Hervé MEMBRE – Monsieur Christophe NEMOS - Madame Natalia DE ISLA
Madame Nathalie MERCIER

66^{ème} section : PHYSIOLOGIE

Monsieur Nguyen TRAN

67^{ème} section : BIOLOGIE DES POPULATIONS ET ÉCOLOGIE

Madame Nadine MUSSE

PROFESSEURS ASSOCIÉS

Médecine Générale

Professeur associé Francis RAPHAEL

MAÎTRES DE CONFÉRENCES ASSOCIÉS

Médecine Générale

Docteur Jean-Louis ADAM
Docteur Paolo DI PATRIZIO
Docteur Sophie SIEGRIST
Docteur Arnaud MASSON

=====

PROFESSEURS ÉMÉRITES

Professeur Jean-Marie ANDRÉ - Professeur Daniel ANTHOINE - Professeur Pierre BEY
Professeur Patrick BOISSEL – Professeur Michel BOULANGÉ – Professeur Jean-Pierre CRANCE
Professeur Jean-Pierre DELAGOUTTE – Professeur Jean-Marie GILGENKRANTZ
Professeur Simone GILGENKRANTZ - Professeur Michèle KESSLER - Professeur Henri LAMBERT
Professeur Denise MONERET-VAUTRIN – Professeur Pierre MONIN - Professeur Jean-Pierre NICOLAS
Professeur Luc PICARD - Professeur Michel PIERSON - Professeur Jacques POUREL
Professeur Jean-François STOLTZ - Professeur Michel STRICKER - Professeur Gilbert THIBAUT
Professeur Hubert UFFHOLTZ – Professeur Paul VERT
Professeur Colette VIDAILHET - Professeur Michel VIDAILHET

DOCTEURS HONORIS CAUSA

Professeur Norman SHUMWAY (1972)
Université de Stanford, Californie (U.S.A)
Professeur Paul MICHELSEN (1979)
Université Catholique, Louvain (Belgique)
Professeur Charles A. BERRY (1982)
Centre de Médecine Préventive, Houston (U.S.A)

Harry J. BUNCKE (1989)
Université de Californie, San Francisco (U.S.A)
Professeur Daniel G. BICHET (2001)
Université de Montréal (Canada)
Professeur Brian BURCHELL (2007)
Université de Dundee (Royaume Uni)

Professeur Ralph GRÄSBECK (1996)
Université d'Helsinki (FINLANDE)
Professeur James STEICHEN (1997)
Université d'Indianapolis (U.S.A)
Professeur Duong Quang TRUNG (1997)
*Centre Universitaire de Formation et de
Perfectionnement des Professionnels de Santé d'Hô
Chi Minh-Ville (VIËTNAM)*
Professeur Marc LEVENSTON (2005)
Institute of Technology, Atlanta (USA)

Professeur Pierre-Marie GALETTI (1982)
Brown University, Providence (U.S.A)
Professeur Mamish Nisbet MUNRO (1982)
Massachusetts Institute of Technology (U.S.A)

Professeur Théodore H. SCHIEBLER (1989)
Institut d'Anatomie de Würzburg (R.F.A)
Professeur Maria DELIVORIA-PAPADOPOULOS
(1996)
Université de Pennsylvanie (U.S.A)
Professeur Mashaki KASHIWARA (1996)
*Research Institute for Mathematical Sciences de Kyoto
(JAPON)*

Professeur Mildred T. STAHLMAN (1982)
Vanderbilt University, Nashville (U.S.A)

REMERCIEMENTS

A notre Maître et Président de thèse,

Monsieur le Professeur Bernard KABUTH

Professeur de Pédopsychiatrie

Docteur en Psychologie

Nous vous remercions infiniment de l'honneur que vous nous faites en acceptant de présider ce jury.

Nous sommes honoré de l'intérêt que vous avez porté à cette thèse tout au long de son élaboration. Nous vous remercions de nous avoir soutenu et rassuré durant ce travail et durant notre parcours d'étudiant cheminant vers les responsabilités. Nous vous remercions de nous avoir accueilli dans votre service et de la confiance que vous avez su placer en nous. Votre disponibilité, votre enthousiasme et vos qualités relationnelles sont un modèle pour notre futur exercice.

Que ce modeste travail soit pour vous, l'expression de notre grande estime, de notre profonde admiration et de notre vive gratitude.

A notre Maître et Juge,

Madame le Professeur Colette VIDAILHET

Professeur Emérite de Pédopsychiatrie

Nous vous sommes très reconnaissant de nous faire l'honneur de juger ce travail.

Nous regrettons de n'avoir pu bénéficier de vos enseignements.

Par ce modeste travail, veuillez recevoir un signe du profond respect et de la gratitude que nous vous témoignons.

A notre Maître et Juge,

Monsieur le Professeur Daniel SIBERTIN-BLANC

Professeur de Pédopsychiatrie

Nous vous remercions d'avoir bien voulu nous faire l'honneur d'accepter de siéger à ce jury.

Nous sommes fier d'avoir été l'un de vos élèves et vous remercions pour la richesse de vos enseignements. Votre volonté de transmettre votre savoir et votre expérience, nous ont porté par la passion que vous y mettez.

Par ce modeste travail, veuillez recevoir un signe de la vive gratitude, de la grande estime et du profond respect que nous vous témoignons.

A notre Juge et Directrice,

Mme le Docteur Claire OBACZ

Docteur en Psychiatrie de l'Adulte

Nous vous remercions infiniment d'avoir accepté de siéger à ce jury et d'avoir diriger notre travail.

Nous sommes honoré d'avoir pu modestement collaborer avec vous pour l'élaboration d'un outil de recherche. Votre motivation et votre rigueur sont des exemples qui devraient nous inspirer longtemps.

Plus personnellement, nous vous remercions pour le soutien et l'empathie que vous nous avez réservés depuis nos premières années d'interne.

Par ce modeste travail, veuillez recevoir un signe de la grande estime et du profond respect que nous vous témoignons.

A tous ceux qui ont participé à ma formation :

A tous mes Professeurs et enseignants de Médecine de Bordeaux, Nancy, et Strasbourg.

A tous mes Maîtres de stages d'internat qui m'ont enseigné la pratique clinique et m'ont offert leur confiance et leur encadrement à Briey, Sarreguemines, Jury, Pont-à-Mousson, Laxou, Mirecourt, Nancy-Brabois et Metz.

A toutes les équipes soignantes qui m'ont accueilli pendant mes stages d'interne et en particulier celles de l'UPPAM et de l'UIAO.

A Françoise, pour le réconfort que tu m'as offert et pour te soucier de l'accueil des internes.

A l'équipe des Consultations Pédiopsychiatriques de Brabois. Vous avez su m'encourager et me faire bénéficier de votre expérience.

A Marie-Chantal, ton aide m'a été plus que précieuse.

A Denis.

A l'équipe du SMPR de la Maison d'Arrêt de Metz Queuleu, pour votre soutien, votre compréhension et votre tendresse. Merci de m'accorder votre confiance, vous croyez en moi et me permettez d'avancer. Je suis très heureux de travailler avec vous.

Au Docteur Horrach, qui m'offre les meilleures conditions d'écoute et de soutien. Vos attentions me touchent particulièrement.

A tous ceux qui ont permis la réalisation de ce travail,

Aux praticiens hospitaliers des secteurs de psychiatrie et de pédopsychiatrie du CPN, en remerciement de nous avoir autorisé et facilité l'accès aux dossiers.

A toute l'équipe du DIM et des archives du CPN. Votre collaboration a été chaleureuse et efficace.

Aux secrétaires du CPN, pour votre disponibilité et votre contribution à la découverte des dossiers.

Au Dr Nelly Agrinier et à Marie-Line Erpelding, du Service Epidémiologie et Evaluations Cliniques du CHU de Nancy Brabois pour la réalisation de l'étude statistique des données.

Au Dr Fabienne Rouyer – Ligier. Conjuguer les compétences de psychiatre aux outils de la Statistique, était l'étape déterminante de ce travail. Merci d'y avoir grandement collaboré. Vos qualités humaines sont très appréciables. Merci pour votre disponibilité et votre douceur.

A ma famille,

Qui malgré les distances, reste proche dans mon cœur.

A toi Nathalie, ma sœur et « petite » maman, avec mon amour et toute ma gratitude,
Pour ta douceur et ta générosité. Tu as toujours été à mes côtés pour les moments difficiles. Merci d'avoir tant pris soin de moi ! Je te remercie pour ta présence bienveillante et respectueusement discrète. Tu es ma confidente pour mes pires tracas même si j'ai du mal à t'appeler.

A toi Rémi, mon frère taquin, avec mon amour et toute ma gratitude,
Pour ton soutien inconditionnel, ton sourire et ton gros cœur bien caché. Je n'ai jamais su t'exprimer mon amour, mais sache que ta simplicité et ta joie de vivre me manquent. Toi, qui sait faire rimer la sagesse et l'aventure. Merci de m'avoir montré la voie.

A mes parents, avec tout mon amour et ma gratitude,
Merci de m'avoir permis de réaliser ses longues études. Je vous suis très reconnaissant pour les sacrifices que vous avez réalisés. Je suis très heureux que vous m'ayez transmis vos valeurs de respect, d'amour et de tolérance qui sont, aujourd'hui, des outils nécessaires à mon métier. Merci d'être si chaleureux avec moi. Votre amour est ma fierté.

A toi Maman,

Tu m'as appris l'honnêteté et le bonheur de s'occuper de ses proches. Tu as toujours voulu me protéger. Gracieuse et avenante, j'ai toujours admiré le feu de ton cœur et la chaleur que tu mets dans la relation à l'autre. Tu m'as transmis ta prudence et ta sensibilité. Pour ta force et ton courage ; pour toute la tendresse et l'amour que tu m'as donnés. Merci.

A toi Papa,

Tu m'as toujours poussé à me surpasser et tu as éveillé mon esprit critique. Tu m'as appris à faire confiance en la nature humaine. Tu m'incites à préférer l'optimisme et l'action devant les embûches de la vie. A créer les moments propices au bonheur plutôt qu'attendre celui-ci. Ces mots ou tu m'invites à profiter intensément de la chance qu'on a de vivre; sont gravés en moi : « Tu dormiras quand tu seras mort ! ». Pour ta philosophie, ta poésie et tes valeurs; pour ton amour, ton esprit, et ta folie si douce. Merci.

A la mémoire de mes grands-parents, merci pour votre amour inconditionnel, merci de m'avoir transmis la fierté de nos racines espagnoles.

A mes tantes et oncles de France et de Navarre, ainsi qu'à mes cousines et cousins. Pour leur sens de la famille, et les souvenirs de vacances au soleil.

A mes amis,

A Loïc, pour avoir supporté mes humeurs et pour tous les moments vécus depuis notre rencontre.

A Claudine, pour ta sagesse et ton affection. Parce que je peux compter sur ta sincérité tout comme sur ta tendresse.

A Josiane, mon amie si généreuse. Pour ton aide, ton soutien et ton amitié inconditionnelle. Tu es un trésor. Tu m'as donné bien plus que tu n'as jamais reçu. Merci de m'aider à prendre la vie du bon côté.

A toutes nos soirées de rires, de larmes et de travail. A nos voyages inoubliables. A votre cuisine savoureuse, surprenante et généreuse. Que la vie nous prodigue encore ces précieux instants partagés.

Aux copains d'internat, Cathy, Lidiana, Elodie, Marine, Joseph, Olivier, Maxime et Jérôme.

A Sophie, mon amie de galère des années concours de médecine.

A Sandra, que j'espère retrouver un jour.

A Gaudéric, mon ami d'enfance et à toute sa famille.

* * *

SERMENT

"Au moment d'être admis à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité. Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux. Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité. J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences. Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences. Je donnerai mes soins à l'indigent et à quiconque me les demandera. Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis dans l'intimité des personnes, je tairai les secrets qui me sont confiés. Reçu à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs. Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés. J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité.

Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré et méprisé si j'y manque".

SOMMAIRE

PREMIERE PARTIE : THEORIE

I. La recherche et ses méthodes scientifiques

A.	Définition.....	20
B.	Catégories de recherches	21
1.	<i>Recherche empirique, recherche théorique</i>	21
2.	<i>Recherche fondamentale, recherche appliquée</i>	21
C.	Les fondements scientifiques de la recherche	22
1.	<i>Observation systématique, description exhaustive et classification objective</i>	22
2.	<i>La rigueur des concepts, la faillibilité des théories</i>	22
3.	<i>La logique des raisonnements</i>	23
4.	<i>L'évaluation de la probabilité</i>	23
5.	<i>L'élaboration et le test d'hypothèses, le caractère provisoire des conclusions</i>	24
6.	<i>La probité des chercheurs</i>	25
D.	Les différents types d'études épidémiologiques.....	25
1.	<i>Les études expérimentales</i>	25
2.	<i>Les études d'observations</i>	26

II. Les particularités de la recherche en pédopsychiatrie

A.	Entre science et recherche de sens : un bref rappel historique.....	34
B.	Les différentes approches et conceptions	37
C.	La subjectivité.....	38

III. Les échelles et leur validation

A.	Définition et intérêt des échelles en pédopsychiatrie.....	41
B.	Caractéristiques métrologiques.....	43
1.	<i>Sensibilité, spécificité</i>	43
2.	<i>Fiabilité, fidélité</i>	43
3.	<i>Validité</i>	43
4.	<i>Consistance interne</i>	44
C.	Nature et modalité de la mesure.....	45

IV. La fidélité inter-juges

A.	Définition et intérêt	46
B.	Le coefficient Kappa de Cohen [45 à 50]	47
1.	<i>Définition</i>	47
2.	<i>Significativité statistique</i>	50
3.	<i>Remarques générales</i>	52
4.	<i>Interprétation de la valeur du Kappa</i>	52
5.	<i>Utilisation pratique</i>	54

DEUXIEME PARTIE : ETUDE

I. Etude initiale faisant l'objet de l'analyse méthodologique

A.	Description de l'étude, matériel et méthode	56
1.	<i>Présentation de l'étude</i>	56
2.	<i>Population étudiée</i>	57
3.	<i>Données étudiées</i>	57
4.	<i>Méthode de recueil</i>	58
B.	Analyse de la méthodologie de l'étude initiale	59
1.	<i>Démarche de recherche</i>	59
2.	<i>Type d'étude</i>	60
3.	<i>Données étudiées et méthodologie de recueil</i>	62

II. Examen de la reproductibilité des résultats au travers de la fidélité inter juges d'un instrument de mesure complexe

A.	Justification de l'examen	63
B.	Elaboration d'un instrument de mesure complexe (IMEC)	64
1.	<i>Justification</i>	64
2.	<i>Méthodologie de construction</i>	64
3.	<i>Séances de mise au point</i>	65
4.	<i>Aspect de l'IMEC</i>	66
5.	<i>Manuel d'utilisation de l'IMEC</i>	67
C.	Objectifs	74
1.	<i>Objectif principal</i>	74
2.	<i>Objectif secondaire</i>	74
D.	Matériel et méthodes	75
1.	<i>Méthodologie de revue de la littérature</i>	75
2.	<i>Type d'examen</i>	75
3.	<i>Population étudiée</i>	75
4.	<i>Modalités de recueil des données</i>	76
5.	<i>Données recueillies</i>	76
6.	<i>Analyse statistique</i>	76
E.	Résultats	77
1.	<i>Répartition des symptômes selon le juge</i>	77
2.	<i>Mesure de la fidélité du relevé de chaque item entre les juges</i>	82
3.	<i>Mesure de la fidélité du relevé par dimensions</i>	87
4.	<i>Pourcentage de concordance brute</i>	90
F.	Limites de la mesure de la fidélité inter-juges	94
G.	Discussion	96
1.	<i>Interprétation de la fidélité inter-juges</i>	96
2.	<i>Synthèse</i>	103
3.	<i>Modifications de l'IMEC</i>	104
H.	Perspectives	117

Références	121
-------------------------	-----

Annexe	126
---------------------	-----

INTRODUCTION

La recherche en psychiatrie réunit d'importants enjeux qui vont considérablement influencer sur l'avenir de la spécialité.

La connaissance de l'esprit constitue sans doute la nouvelle conquête du XXIème siècle, après celle du génome. Elle concerne l'ensemble des neurosciences au sein desquelles la psychiatrie est désormais intégrée par les uns et ignorée par les autres. Le choix propre de notre discipline par rapport à cet enjeu est évidemment essentiel.

La psychiatrie émerge en effet d'une sorte d'extra-territorialité par rapport au champ médical et social, dans lequel elle se trouve désormais directement impliquée. La place de la subjectivité dans l'investigation clinique et son application dans la recherche ont toujours été au centre de ses préoccupations. Cependant l'intégration entre la méthode scientifique et les particularités de l'objet psychiatrique restent le lieu de débats.

En fait, la mise au premier plan de la nécessité de recherche en psychiatrie et en santé mentale réactualise de façon quelquefois caricaturale les problèmes d'articulation entre la perspective clinique et celle des neurosciences.

Dans une première partie théorique, nous étudierons ce qui fait la force de la méthode scientifique de recherche basée notamment sur l'objectivité et la reproductibilité des résultats. Nous envisagerons également les particularités de la recherche en pédopsychiatrie, par la spécificité de son objet et de son approche et par la place singulière qui tient la subjectivité. Nous définirons l'utilisation des échelles d'évaluation dans cette discipline et présenterons les principes de la fidélité inter-juges comme mesure de la reproductibilité des résultats.

Dans une seconde partie, nous tenterons de mettre en application les éléments théoriques abordés au travers d'une étude de recherche en pédopsychiatrie menée récemment à Nancy. Nous présenterons l'étude initiale qui fait l'objet de notre analyse avant d'examiner la méthodologie employée. Nous investiguerons plus particulièrement la fidélité inter-juges d'un instrument de mesure qui y est employé. Enfin, nous proposerons une discussion sur l'interprétation de ces résultats avant de proposer des modifications de l'instrument et des perspectives de poursuite de l'étude initiale réalisée.

1^{ERE} PARTIE : THEORIE

I. La recherche et ses méthodes scientifiques [1,2]

A. Définition

Selon l'OMS, la recherche consiste « en une quête des connaissances au moyen d'une enquête, d'une étude ou d'une expérimentation conduite avec application dans le but de découvrir et d'interpréter des connaissances nouvelles ».

« La méthode scientifique est un ensemble systématique de procédures et de techniques relatives à l'exécution d'une étude ou d'une expérience destinée à acquérir de nouvelles connaissances ». La recherche médicale peut être considérée comme une suite d'interrogations décisives conduisant à la découverte de faits ou d'informations qui améliorent notre compréhension de la santé et des maladies de l'être humain.

Elle consiste en une démarche rationnelle, organisée et rigoureuse, pour étudier et comprendre. Elle élève le niveau de la pensée, approfondit par la réflexion et la critique des chantiers déjà ouverts, explore par le raisonnement, l'intuition et l'expérience des domaines encore inconnus de notre univers. Une telle démarche comporte, par ailleurs, des moments de création d'où surgit la formulation d'hypothèses et d'approches inédites, qui vont permettre de renouveler les perspectives et les méthodologies, et de procéder à des innovations. La recherche a pour fonction première la formulation de questions nouvelles et la production de nouveaux savoirs; elle contribue à créer ou à baliser le futur par le progrès de tous les domaines de la connaissance, de même que par la diffusion et le partage de ces avancées avec la société.

B. Catégories de recherches

1. Recherche empirique, recherche théorique

L'approche philosophique de la recherche est fondamentalement de deux types : empirique et théorique.

Les recherches dans le domaine de la santé sont principalement de nature empirique, c'est-à-dire qu'elles sont fondées plus sur l'observation et l'expérience que sur la théorie et l'abstraction. La recherche épidémiologique, par exemple, dépend de la collecte systématique d'observations sur des phénomènes de santé spécifiques au sein d'une population définie.

La recherche empirique et la recherche théorique se complètent pour développer la compréhension des phénomènes, pour prévoir des événements futurs, ainsi que pour prévenir des événements nuisibles au bien-être général de la population concernée.

La prise en compte du hasard ou de la probabilité est déterminante dans la recherche; elle est à la base du plan de recherche. Pour en assurer la validité, ce plan de recherche doit tenir compte, avant toute autre chose et en permanence, du rôle du hasard.

Ce sont les méthodes statistiques qui permettent de respecter les lois de probabilité au cours des études, et d'analyser et d'interpréter correctement les résultats. La statistique est l'outil qui permet à la recherche dans le domaine de la santé d'être empirique plutôt qu'abstraite ; elle permet de confirmer les constatations au moyen d'observations et d'expérimentations supplémentaires.

2. Recherche fondamentale, recherche appliquée

Du point de vue fonctionnel, la recherche peut être fondamentale (ou pure) ou appliquée.

La **recherche fondamentale** se consacre à l'acquisition des connaissances sans but défini d'utilité ou d'objet spécifique. En psychiatrie, elle s'organise autour de trois orientations : la pharmacologie, la génétique et la psychophysiologie.

La **recherche appliquée** découle d'un problème et elle est dirigée pour résoudre un problème existant. Il s'agit d'une recherche clinique qui porte sur des personnes totales et sur la situation où elles sont placées. Par exemple, la recherche sur les conduites suicidaires ou sur un type particulier de pathologie comme la schizophrénie, la dépression, ...

C. Les fondements scientifiques de la recherche

La force de la méthode scientifique est décrite par nombre d'auteurs [3,4] et repose sur les caractéristiques que nous allons décrire :

1. Observation systématique, description exhaustive et classification objective

En science, les observations ne doivent rien laisser au hasard et se dérouler de manière ordonnée et complète. Elles doivent être le fait de nombreux observateurs indépendants les uns des autres.

Les descriptions doivent être systématiques et exhaustives. Leur classification doit se faire sur une base logique et objective.

L'objectivité est la capacité de reconnaître ce qui est réel, indépendamment de tout préjugé et de toute interprétation personnelle. C'est le contraire de la subjectivité. Etre objectif, c'est demeurer neutre, impartial devant les faits.

2. La rigueur des concepts, la faillibilité des théories

Les expériences doivent être menées et décrites de manière rigoureuse de façon à ce que le plus haut degré possible d'objectivité soit atteint. Toute expérience doit être décrite de manière à pouvoir être reproduite et vérifiée.

Les concepts utilisés doivent également être définis de manière rigoureuse, sans aucune équivoque, ou à tout le moins de la manière la plus exacte possible.

Les théories qui expliquent les phénomènes en faisant intervenir des facteurs ou des causes particulières doivent être faillibles, c'est-à-dire qu'elles doivent avoir des conséquences concrètes qu'on peut potentiellement observer et qui vont permettre de confirmer ou d'infirmer les théories en question.

3. La logique des raisonnements

Le raisonnement, ou l'inférence, est la force qui permet les avancées dans la recherche.

Qu'il soit inductif ou déductif, il doit être parfaitement clair et doit obéir aux lois de la logique.

Cela signifie qu'un énoncé ou une conclusion doit être accepté parce qu'un ou plusieurs autres énoncés ou prémisses (évidences) sont vrais. Des suppositions, des présomptions ou des théories peuvent être ainsi développées par inférence selon une construction minutieuse, en vue de postuler des hypothèses à tester. Le test d'hypothèses est la méthode de base pour faire avancer les connaissances dans les sciences et de surcroît en psychiatrie.

Dans la déduction, la conclusion résulte nécessairement des prémisses, comme dans le syllogisme (tout A est B, tout B est C, donc tout A est C) ou les équations algébriques.

La déduction se distingue par le fait qu'elle va du général au particulier et qu'elle ne tolère aucun élément de hasard ou d'incertitude. Les inférences déductives conviennent donc tout particulièrement à la **recherche théorique**.

Dans l'induction, la conclusion ne résulte pas nécessairement des prémisses ou de l'évidence (faits). On peut dire seulement que la conclusion est plus probablement valide si les prémisses sont vraies, c'est-à-dire, qu'il y a une possibilité d'avoir les prémisses vraies mais la conclusion fausse. Il faut donc tenir compte de tous les effets du hasard. Par ailleurs le raisonnement inductif se distingue par le fait qu'il va du particulier vers le général, donc il construit. Étant essentiellement empirique, **la recherche en psychiatrie, comme toute recherche en matière de santé, repose presque entièrement sur des raisonnements par induction.**

4. L'évaluation de la probabilité

Le raisonnement scientifique utilise les lois de la probabilité, c'est un **raisonnement statistique**.

La probabilité est une mesure de l'incertitude ou de la variabilité des caractéristiques parmi les

individus d'une population. Si l'on observe une population entière, le calcul des fréquences relatives des variables fournit toute l'information sur la variabilité. Si nous observons seulement un échantillon d'individus tiré de la population, l'inférence de l'échantillon vers la population (extrapolation du particulier au général), impliquera l'identification des probabilités des événements en cours d'observation, ainsi que le recours aux lois de probabilité qui nous permettent de mesurer le degré d'incertitude dans nos inférences.

En se basant sur les probabilités calculées, on accepte ou on rejette l'hypothèse ; ou bien l'état d'incertitude reste non résolu, ce qui est le cas particulier lorsque la taille de l'échantillon est trop petite pour assurer la fiabilité.

L'exigence décisive pour assurer la validité du plan de recherche, est l'évaluation de la probabilité du début à la fin.

Les éléments du plan les plus notables qui doivent assurer l'intégrité de la probabilité et la prévention de biais, sont les suivants : échantillonnage représentatif, randomisation dans la sélection des groupes d'étude, maintien de groupes de comparaison servant de témoins, expériences et sujets traités en double aveugle, et utilisation de méthodes statistiques de probabilité dans l'analyse et l'interprétation des résultats.

5. L'élaboration et le test d'hypothèses, le caractère provisoire des conclusions

Les hypothèses sont des énoncés soigneusement construits relatifs à un phénomène dans la population. L'un des outils les plus utiles dans les recherches sur la santé est la génération d'hypothèses qui, après avoir été testées, conduiront à l'identification des causes les plus probables d'une maladie. Bien que nous ne puissions tirer des conclusions définitives ou revendiquer la preuve en utilisant des méthodes inductives, nous pouvons nous approcher toujours plus près de la vérité en réfutant des hypothèses existantes et en les remplaçant par d'autres plus plausibles.

Enfin, toute conclusion scientifique, aussi ferme soit-elle, doit être considérée comme vraie jusqu'à ce qu'on ait prouvé le contraire ou qu'on dispose d'une meilleure explication, plus exacte ou plus complète. Il n'y a pas de vérité définitive ni de vérité absolue en science.

6. La probité des chercheurs

Quels que soient les résultats — que l'hypothèse de recherche soit confirmée, infirmée ou reformulée —, toute recherche scientifique aboutit à l'obligation d'entreprendre de nouvelles recherches.

Les chercheurs ne doivent jamais fausser ou modifier leurs résultats afin de les rendre conformes à leurs théories ou à leurs hypothèses. Ils doivent se soumettre volontiers à toute vérification extérieure au besoin.

Ainsi, la recherche scientifique est-elle par définition un cycle sans fin, chaque recherche suscitant un ou plusieurs questionnements nouveaux. Ce processus est indépendant des individus particuliers. Il se peut qu'une recherche donne naissance à un nouveau sujet de recherche qui sera étudié bien des années plus tard ou à des milliers de kilomètres du lieu où la recherche a débuté. Ce caractère « impersonnel » de la recherche scientifique lui permet de progresser grâce à l'universalité de ses méthodes à travers des générations de chercheurs, indépendamment des frontières.

D. Les différents types d'études épidémiologiques [5 à 17]

La recherche repose sur deux grands types d'études épidémiologiques : les études **expérimentales** et les études d'**observations** (descriptives ou analytiques).

1. Les études expérimentales

La meilleure vérification des hypothèses est réalisée par des expériences dans lesquelles tous les facteurs, sauf celui en considération, peuvent être maîtrisés. Cependant, pour des raisons éthiques et pratiques, cette méthode est rarement possible dans le contexte des maladies humaines. On procède alors plutôt par des expériences dites « naturelles » ou des observations soigneusement planifiées comportant suffisamment d'informations sur les facteurs « parasites » pour permettre de corriger leurs effets lors des inférences.

2. Les études d'observations

Dans ces méthodes, on compare des groupes d'individus pour relever les différences en matière d'exposition ou de résultat. Elles se distinguent des expériences par le fait qu'il n'y a pas d'intervention directe du chercheur qui ne peut maîtriser aucun des facteurs tiers (ou parasites) agissant sur l'un ou l'autre des individus en observation.

Les études basées sur l'observation emploient généralement la méthode d'enquête sur échantillon, dans laquelle on observe différentes caractéristiques d'un échantillon de la population.

On peut y procéder en interrogeant effectivement le sujet, en obtenant les mesures de caractéristiques physiques, ou simplement en extrayant des informations de sources existantes telles que les registres de déclaration des maladies, les dossiers médicaux des hôpitaux ou des dossiers des services d'emploi. Elles peuvent être **descriptives** ou **analytiques**.

a) *Les études descriptives*

(1) Définition

Le trait caractéristique de cette catégorie d'études est qu'elle s'attache avant tout à décrire une situation plutôt qu'à tester une hypothèse ou à rechercher une causalité.

L'étude **descriptive** n'a pas la structure formelle d'une étude analytique ou d'une étude expérimentale, autrement dit, elle n'a pas pour objet spécifique de tester une hypothèse ou rechercher une causalité, mais de décrire une situation. La richesse des informations obtenue dans la plupart des études descriptives permet la génération d'hypothèses que l'on peut ensuite tester par un plan d'étude analytique ou expérimentale. Cependant, la démarche descriptive peut être intégrée ou ajoutée aux méthodes utilisées dans ces deux derniers cas car elle permet d'**enrichir considérablement la base de données**.

Une étude descriptive nécessite la collecte, l'analyse et l'interprétation de données. On peut faire appel à des techniques qualitatives et quantitatives à l'aide de questionnaires, d'interrogatoires, des observations de participants, de statistiques des services, ainsi que de documents donnant des renseignements sur des communautés, des groupes, des situations, des programmes et d'autres unités individuelles ou écologiques.

(2) Types d'études descriptives

(a) *Séries de cas*

Ce type d'étude est basé sur les comptes-rendus d'une série de cas d'une pathologie donnée ou d'une série de cas traités, sans affectation spécifique d'un groupe témoin. Dans ce type d'étude on ne considère que le nombre de cas de la maladie, qui représente le numérateur, si bien qu'il ne doit pas servir pour faire une évaluation de risques.

(b) *Diagnostic ou évaluation des besoins d'une communauté*

Ce type d'étude nécessite la collecte de données sur la situation actuelle concernant les problèmes sanitaires, programmes de santé, progrès réalisés, contraintes, couches sociales, forces vives, noyaux de résistance ou de forte prévalence ou groupes particulièrement exposés. Son but est d'identifier les besoins et de fournir les données de base qui orienteront les études complémentaires ou les actions à mener.

(c) *Description épidémiologique des cas de maladies*

On adopte souvent l'approche descriptive à cette fin. Il faut recueillir des données sur le nombre de cas et sur la répartition de la maladie dans des populations en fonction de certaines caractéristiques spécifiques des individus (par exemple, âge, sexe, niveau d'instruction, consommation de tabac, religion, profession, catégorie sociale, état civil, santé, personnalité) ; en fonction du lieu (rural ou urbain, local, régional, national, international) ; et en fonction du facteur temporel (épidémique, saisonnier, cyclique, séculaire). Une description peut porter aussi sur des caractéristiques familiales : rang de naissance, nombre d'enfants, taille de la famille, âge maternel, espacement des naissances, composition de la famille.

(d) *Enquêtes transversales descriptives ou enquêtes de population*

Comme son nom l'indique, une enquête transversale nécessite la collecte de données sur une coupe transversale d'une population, qui peut être la population totale ou une partie (échantillon) de la population. Beaucoup d'enquêtes transversales ne cherchent pas à tester une hypothèse

concernant une association, et appartiennent donc à la catégorie des études descriptives. Elles fournissent un taux de prévalence à un moment donné dans le temps (prévalence ponctuelle) ou sur une période de temps (prévalence périodique). La population à risque, objet de l'étude, sert de dénominateur pour le calcul des taux de prévalence.

Dans ce type d'étude descriptive on retrouve les enquêtes d'évaluation de la distribution d'une maladie, d'une incapacité, d'un état pathologique, d'une condition immunologique, d'un état nutritionnel, de l'état physique, de l'intelligence et ainsi de suite. On peut suivre le même plan d'étude dans la recherche sur les systèmes de santé pour décrire 'la prévalence' par certaines caractéristiques – la tendance d'utilisation des services de santé et de l'observance – ou dans des sondages d'opinion.

(e) Etudes descriptives écologiques

Lorsque l'unité d'observation est un agrégat (par exemple, une famille, un clan ou une école) ou une unité écologique (un village, une ville ou un pays) l'étude devient une étude descriptive écologique.

Comme nous l'avons déjà souligné, en règle générale, le test d'hypothèse n'est pas un objectif de l'étude descriptive. Cependant, il peut trouver sa place dans certaines études descriptives citées plus haut (enquêtes transversales, études écologiques).

D'autre part, la description des données fait partie intégrante de l'étude analytique. Par conséquent, l'étude descriptive peut représenter un premier temps dans une étude analytique.

b) Les études analytiques

(1) Définition

Les études analytiques ont pour principal objectif d'établir la relation (association) entre un "facteur de risque" (agent étiologique) et un résultat (maladie). Dans ce type d'étude de causalité, **le test d'hypothèse est le premier outil d'inférence**. La démarche essentielle dans une étude

analytique est d'élaborer une hypothèse spécifique susceptible d'être mise à l'épreuve et de faire un plan d'étude qui permet de maîtriser les variables parasites éventuelles qui risqueraient de confondre la relation observée entre le facteur étudié et la maladie. Cette démarche varie en fonction de la stratégie analytique adoptée : les études longitudinales (cas-témoins ou de cohorte) et les études transversales.

(2) Types d'études analytiques

(a) *Etudes cas-témoins*

En épidémiologie, la stratégie analytique la plus simple et la plus souvent adoptée fait appel à l'étude comparative cas-témoins. Essentiellement, elle consiste à établir les causes de maladies en recherchant les relations ou associations entre l'exposition à un facteur de risque et l'apparition de la maladie.

Son plan est relativement simple et se base toujours sur **les antécédents des cas et des témoins** en matière d'exposition au risque.

L'étude cas-témoins est donc toujours **rétrospective**.

Dans ce type d'étude, on recherche une association en comparant l'exposition d'une série de cas de la maladie concernée avec la situation d'exposition de groupes témoins soigneusement sélectionnés qui sont exempts de cette maladie.

L'analyse des données permet de déterminer si les conditions d'exposition des cas étaient différentes de celles des témoins. Le facteur de risque s'est produit ou a commencé dans le passé, probablement avant l'apparition de la maladie, par exemple, l'habitude de fumer ou une infection antérieure ou la prise d'un médicament. On recueille les données relatives à l'exposition en se renseignant sur les antécédents et/ou en consultant les dossiers médicaux. Parfois le facteur ou l'attribut incriminé est une condition permanente, comme le groupe sanguin, que l'on détermine par un examen clinique ou de laboratoire. Une fréquence plus élevée de l'attribut ou du facteur de risque chez les cas de maladie que chez les témoins indique une relation entre celui-ci et la maladie.

Cette relation peut avoir une signification étiologique. Autrement dit, si la proportion des cas de maladie ayant des antécédents en matière d'exposition, ou des dossiers indiquant une exposition par le passé, est plus élevée que celle des témoins, on peut considérer que le facteur ou l'attribut est un facteur causal.

Avantages des enquêtes cas-témoins

Voici quelques avantages des études cas-témoins :

- elles sont réalisables plus rapidement et peu coûteuses, elles peuvent être facilement répétées ;
- elles sont applicables lorsque la maladie à étudier est rare ;
- elles sont relativement efficaces, avec un échantillon plus petit que dans le cas d'une étude de cohorte ;
- elles posent peu de problèmes de perte de participation (attrition), comme dans un suivi avec examens périodiques où certains sujets refusent de coopérer ;
- elles offrent la possibilité d'examiner plusieurs facteurs de risque d'une même maladie ;
- c'est parfois la stratégie par observation la plus précoce praticable pour déterminer une association.

Inconvénients et sources de biais des enquêtes cas-témoins

Voici certains des problèmes liés aux études cas-témoins :

- puisqu'il n'y a pas de dénominateur épidémiologique commun (la population à risque) tout calcul des taux d'incidence, et donc des risques attribuables, est impossible ;
- la séquence temporelle est un problème sérieux pour beaucoup d'études cas-témoins où il est impossible de savoir si l'attribut a précédé la maladie/état pathologique ou en est le résultat ;
- la sélection des cas étudiés et des témoins est très sujette au biais. Ceci est particulièrement grave lorsqu'un seul groupe témoin est associé au facteur de risque étudié ;
- il peut être très difficile voire impossible d'obtenir des informations sur l'exposition si la période de remémoration est longue ;
- la survie sélective, qui a un effet dans les études cas - témoins, peut biaiser la comparaison. Il n'y a aucun moyen de savoir si l'exposition était identique pour ceux qui sont morts et ceux qui ont survécu ;
- un biais de mesures est possible, y compris par la remémoration sélective ou l'erreur de classification (des cas classés avec les témoins, et réciproquement). Il existe aussi l'effet de Hawthorne, selon lequel, la répétition des interrogatoires, influence les réponses données par les sujets étudiés ;

- les études cas-témoins sont incapables de révéler d'autres états liés au facteur de risque : par exemple, dans une étude des effets secondaires des contraceptifs oraux, il faut connaître d'avance ces effets secondaires pour pouvoir établir un plan d'étude cas-témoins.
- enfin, les biais sont difficilement mesurables.

(b) Etude de cohorte (enquête exposés/non exposés)

La stratégie classique des études de cohorte est de prendre au départ une population de référence (ou un échantillon représentatif de celle-ci), dont certains individus présentent des caractéristiques ou des attributs d'intérêt pour l'étude (groupe exposé), tandis que d'autres individus ne les présentent pas (groupe non exposé). Les deux groupes doivent être exempts de la pathologie en considération au commencement de l'étude. On part des expositions et on recherche les malades. Il s'agit donc d'études **prospectives**. L'observation des deux groupes se poursuit sur une période donnée afin de découvrir pour chaque groupe son risque de développer la pathologie objet de l'étude. Les études de cohorte historiques, sont un cas particulier de cohorte car elles sont rétrospectives : l'inclusion de la population s'effectue dans le passé puis elle est suivie pour relever si apparaît la maladie.

Avantages des études de cohorte

Parmi les avantages d'une étude de cohorte par rapport à une étude cas-témoins, on cite :

- comme la population à risque est bien définie, les études de cohorte permettent une mesure directe du risque relatif de développer l'état pathologique dans la population exposée au facteur de risque par rapport à la population non exposée à ce facteur, grâce aux mesures d'incidence calculées séparément pour chaque population ;
- dans une étude de cohorte, on sait que la caractéristique ou facteur de risque précède l'apparition de la maladie, puisque tous les sujets sont exempts de la maladie au commencement de l'étude. Ceci est une condition nécessaire (mais non suffisante) pour conclure à une relation de cause à effet ;
- étant donné que la présence ou l'absence du facteur de risque est constatée avant l'apparition de la maladie, aucun biais n'est introduit par le fait que le sujet apprend son état pathologique, comme cela se produit dans les études cas-témoins ;
- il existe également un moindre risque de problèmes de survie sélective ou de remémoration sélective, même si un biais de sélection n'est pas totalement écarté, puisque certains sujets qui

tombent malades au début de l'étude ne sont plus pris en considération ;

- les études de cohorte peuvent identifier d'autres maladies éventuellement associées au même facteur de risque ;
- contrairement aux études cas-témoins, les études de cohorte permettent d'estimer des risques attribuables, indiquant ainsi la part absolue de la maladie attribuable au facteur de risque ;
- enfin, en prenant un échantillon respectant les règles de probabilité dans la population de référence, il est possible d'extrapoler de l'échantillon à la population de référence avec un degré de précision connue.

Inconvénients des études de cohorte

Parmi les inconvénients des études de cohorte on peut citer :

- ces études sont de longue durée ce qui pose un **problème de faisabilité** ; elles sont relativement **peu efficaces pour l'étude de pathologies rares** ;
- elles sont exigeantes en temps, en personnel, en locaux, et en planification de suivi des patients, et sont donc très coûteuses ;
- la taille des échantillons requise pour une étude de cohorte est très grande ; en particulier pour des pathologies rares, il est difficile de constituer et de gérer des échantillons de la taille nécessaire ;
- le problème le plus grave est celui de l'attrition, ou la perte d'effectif dans les groupes exposés et non exposés au cours de la durée de l'étude par suite de migration ou de refus de continuer de participer à l'étude. L'attrition peut entacher la validité des résultats, si elle a pour effet de rendre les échantillons moins représentatifs ou si les patients qui participent au suivi sont différents de ceux qui ont abandonné. Plus la perte d'effectif est élevée (supérieure à 10-15 %) plus grand sera le risque de résultats biaisés. Le risque d'attrition existe aussi pour les chercheurs, qui peuvent se désintéresser de l'étude, changer d'activité, ou se consacrer à un autre projet de recherche ;
- au cours d'une longue période, on assiste à beaucoup de changements de l'environnement, chez les personnes comme dans la nature de l'intervention, et ceux-ci peuvent créer une confusion en ce qui concerne l'association ou le risque attribuable ;
- au cours d'une longue période, le protocole de l'étude peut influencer le comportement des

patients étudiés de façon à modifier l'évolution de la maladie (effet Hawthorne). Cet effet se produit surtout dans les études où il y a beaucoup de contacts avec les participants, par exemple dans les études de régime ou de contraception. Les changements de comportements posent aussi un problème sérieux dans les sondages d'opinion, les études d'acceptabilité et les enquêtes de nature psychologique.

- enfin, un problème éthique très préoccupant peut se poser si l'on constate que le surnombre de cas de maladies de la population exposée est très significatif avant la fin de la période de suivi.

Même si une étude de cohorte est proche de l'essai randomisé (expérience) du point de vue de sa puissance épidémiologique, il faut souligner la possible existence de problèmes de validité. Il faut prendre soin de s'assurer qu'elle satisfait aux autres exigences de la recherche épidémiologique, en particulier pour ce qui concerne le choix de l'échantillon, la constitution des groupes de comparaison, le traitement des données manquantes, la mise en oeuvre de méthodes statistiques appropriées, et d'autres exigences de base d'un plan de recherche solide.

(c) Enquête analytique transversale

Les études du type transversal, où l'on collecte simultanément les données de cause et d'effet, sans pouvoir déterminer la séquence temporelle, sont considérées comme génératrices d'hypothèses.

Dans une enquête analytique transversale, le chercheur mesure l'exposition et la maladie simultanément sur un échantillon représentatif de la population.

En prenant un échantillon représentatif, il est possible d'extrapoler les résultats obtenus pour l'échantillon à la population dans son ensemble. Les études transversales mesurent l'association entre la variable "exposition" et la maladie présente (prévalence), alors que les études de cohorte mesurent le taux de progression de la maladie (incidence).

Les maladies rares, les états pathologiques de courte durée, les maladies à létalité élevée, peuvent échapper à l'image instantanée saisie par l'enquête transversale. Ces enquêtes transversales sont mieux adaptées à la mesure de la relation entre des caractéristiques semi-permanentes des individus et des maladies chroniques ou états pathologiques stables.

Avantages de l'enquête transversale

Parmi les avantages des enquêtes transversales on peut citer :

- le premier avantage des enquêtes transversales comparées aux études cas-témoins est le fait qu'elles disposent au départ d'une population de référence d'où elles prélèvent les cas et les témoins ;
- leur durée peut être plus courte, donc elles sont moins coûteuses que des études prospectives ;
- elles constituent la première étape d'une étude de cohorte prospective pour dépister et éliminer les cas présentant déjà l'état pathologique ;
- elles fournissent une grande quantité de données très utiles pour la recherche sur les systèmes de santé ;
- elles permettent de formuler un énoncé du risque mais qui manque de précision.

Inconvénients de l'enquête transversale

Les études transversales posent les difficultés suivantes :

- elles ne permettent pas une estimation directe du risque ;
- elles sont sujettes au biais lié à la survie sélective ;
- étant donné que l'exposition et la maladie sont mesurées au même moment au cours du temps, il n'est pas possible de déterminer la séquence temporelle (c'est-à-dire si l'exposition ou la présence d'une caractéristique précède l'apparition de la maladie).

II. Les particularités de la recherche en pédopsychiatrie

A. Entre science et recherche de sens : un bref rappel historique

La formation médicale, l'assise biologique de la médecine, le souci d'une classification universelle des maladies et le développement de la médecine fondée sur les preuves ou médecine factuelle (l'evidence-based medicine, EBM) tirent la médecine en direction d'une science.

Mais la médecine est également une discipline interprétative : elle travaille à interpréter les signes d'un corps malade. La constellation des signes cliniques qui investissent le corps malade impose

au médecin de retrouver une histoire sous la généralité d'un cas répertorié dans la nosographie. La consultation médicale, dans son colloque singulier, tend à parler d'art médical.

Ces deux lectures sont reprises, souvent scolairement, dans l'opposition des sciences de la nature explicatives et des sciences humaines compréhensives.

En psychiatrie, et plus particulièrement en psychiatrie de l'enfant et de l'adolescent, ces deux aspects sont indissociables. La discipline s'appuie à la fois sur les sciences de la nature dites "dures" et sur la nécessité de comprendre aussi bien que possible les anomalies psychiques et la personne qui en souffre. La recherche de sens apparaît indispensable à l'étude des troubles psychiques. Les signes n'ont pas de valeur en soi mais en ce qu'ils parlent du sujet. Ils sont à comprendre dans l'histoire du patient.

Il s'agit non seulement de déterminer la cause d'un trouble psychique mais également de lui trouver une signification. C'est à dire qu'il s'agit d'expliquer mais aussi de comprendre [18].

La méthode scientifique s'est développée avec les sciences naturelles sur une période très longue. Son application dans le champ des sciences humaines n'est pas aisée.

En effet, la compréhension de l'être humain semble être un phénomène plus subjectif, s'appuyant sur des capacités d'empathie qui varient d'une personne à l'autre ou d'une culture à l'autre, alors que l'observation des sciences naturelles est objective, avec des résultats qui doivent être identiques pour tous.

D'illustres auteurs ont souligné cette problématique :

Freud a souligné que certains états mentaux et certains comportements apparemment dénués de sens peuvent être compris comme pourvus de signification, et qu'une intervention dans le processus de signification pouvait réaliser un changement et ainsi devenir causal [19].

Freud, le neurologue, a reconnu que si tel était le cas, alors les processus mentaux significatifs devaient avoir une inscription dans les processus cérébraux. Mais Freud reconnaissait également que certaines questions ne pouvaient alors avoir de réponse, compte tenu de l'état des sciences cognitives : ces questions concernent l'architecture et les caractéristiques fonctionnelles du cerveau concernant une telle inscription.

Jaspers, psychiatre et philosophe allemand, fut un des premiers à saisir la portée de cette problématique de la psychiatrie, dès le début du XXème, et soutenait de manière égale les deux méthodologies [20]: l'importance à la fois des sciences de la psychopathologie et le besoin indispensable de comprendre la question du sens par l'empathie.

Une telle intégration, entre la recherche de sens et la science n'interviendra qu'un demi-siècle plus tard, et les décennies suivantes furent consacrées à une approche plus primitive du conflit, et notamment un clivage opposant causalité et signification, opposant explication et compréhension, sciences comportementales à herméneutique non scientifique. Ces clivages ont connu leur phase la plus violente dans les années 1960, avec des attaques véhémentes continues contre Freud, l'accusant de ne pas être scientifique mais uniquement préoccupé de signification, et des attaques également féroces contre la psychiatrie dans son ensemble, accusée de vider la folie de sa signification, de la déshumaniser même de façon scientifique.

Un signe majeur de l'intégration entre méthodologie scientifique et phénomène de l'esprit, et signification, commence à apparaître dans les années 1960, avec ce que l'on appelle la révolution cognitive en psychiatrie [21].

Depuis lors, les paramètres de la cognition se situent au centre des meilleurs travaux dont nous disposons dans le domaine des sciences du comportement

Des méthodes permettant de mesurer les significations voient le jour dans les années 1980. On retrouve des instruments mesurant des significations et utilisés pour identifier des processus causaux, tels que :

- *the Life Event and Difficulty Schedule* [22] créé par George W. Brown qui propose un modèle psychosocial d'agents précipitants et de facteurs de vulnérabilité liés à la signification et à l'impact d'un événement, en fonction du contexte factuel dans lequel il se produit et en fonction de la biographie.

- *the Adult Attachment Interview Schedule* [23] et le travail de Main et Goldwyn qui montre, entre autre, que les parents ayant eu un attachement insécure dans leur enfance tendent à reporter des attentes idéalisées sur les expériences d'attachement avec leurs enfants.

Les significations personnelles, familiales et sociales, si elles se situent dans le registre herméneutique par opposition aux sciences causales, n'en sont pas pour autant hors d'atteinte de l'investigation et de la théorie scientifique. Le sens peut réguler le comportement de manière causale, et donc tout modèle de déterminisme du comportement qui omet ce fait restera inadéquat. Cela pose certaines difficultés méthodologiques car la question du sens investigate la relation c'est à dire qu'elle dépend à la fois de l'objet observé, de l'observateur et des interactions entre les deux.

La question de l'intégration réelle entre les deux approches, celle du sens et celle de la science, a donc fait l'objet de nombreuses controverses et demeure une question d'actualité en psychiatrie.

B. Les différentes approches et conceptions

S'il est nécessaire de placer la psychiatrie dans le champ scientifique, la tâche se montre ardue en raison de l'ambiguïté liée à la question permanente de la relation entre normalité et pathologie. Les limites entre le normal et le pathologique sont floues.

L'objet de la connaissance psychiatrique est difficilement saisissable et il n'apparaît possible, non seulement d'en rendre compte, mais même de le désigner qu'en multipliant les plans du cadre de référence. Cette difficulté est, en effet, liée à sa nature propre et il y a tout lieu de penser que le cadre général de la psychiatrie doit admettre des modèles différents, situés sur des plans multiples, qui ont chacun des degrés de pertinence variant selon le cas étudié [24].

Quatre modèles coexistent ainsi dans la pratique clinique de la psychiatrie [25] :

- un modèle nosographique : catégoriel ou dimensionnel, qui est particulièrement compatible avec le modèle de recherche biomédical organiciste ; il implique une réaction thérapeutique construite sur la logique diagnostic-traitement ;
- un modèle phénoménologique où entre en compte le vécu du patient, c'est-à-dire l'interaction entre le sujet et son environnement, tel que le sujet l'expérimente en lui-même ;

- un modèle psychanalytique qui, partageant avec le modèle précédent l'intérêt porté à l'intériorité, le canalise selon un *défilé de concepts opératoires* qui s'intéressent à l'inconscient, à la signification cachée des phénomènes psychiques.
- enfin, le modèle systémique où l'accent est porté sur les interactions de communication entre les sujets de la relation.

Dans la psychiatrie française, la volonté de prendre en compte la multiplicité des points de vue et des modèles a pris le pas sur les tendances à l'hégémonie d'une théorie sur toutes les autres. Malgré la forte pression du modèle biomédical devenu économiquement dominant et les contraintes évaluatives qui imposent de privilégier le dénombrable sur toute autre considération, la psychiatrie française se caractérise, par la place qu'y garde le débat théorique et la possibilité même de penser certains phénomènes subjectifs devenus négligeables pour une bonne partie de la psychiatrie internationale [25].

En pédopsychiatrie, les limites entre le normal et le pathologique sont encore plus floues et la distinction entre les troubles est définie moins clairement que chez l'adulte [26].

En conséquence, les difficultés méthodologiques concernant la recherche scientifique s'y trouvent accrues.

C. La subjectivité

La question de l'interprétation, l'opposition entre objectivité et subjectivité est au cœur de débats en médecine. L'Evidence Based Medicine (EBM), courant dominant dans la médecine contemporaine, entend éliminer les aléas découlant des opérations herméneutiques subjectives, grâce à la collecte massive des faits médicaux et à leur traitement statistique objectif. Les actions thérapeutiques devraient être déterminées par les indications fournies par des données élaborées par la statistique.

Cependant, le monde réel en psychiatrie n'est-il pas bio-psycho-social-subjectif comme le souligne J.S Strauss [27] ? Même si on ne peut pas encore tout expliquer, nous ne devons pas exclure la subjectivité de notre modèle de la maladie mentale.

En psychiatrie, encore plus que dans les autres spécialités médicales, la notion de subjectivité est essentielle, du fait de la nature des moyens dont les cliniciens disposent pour procéder au diagnostic.

Il existe bien peu de signes objectifs en psychiatrie, car la plupart des signes qui constituent le tableau clinique se déduisent à partir de l'entretien durant lequel la subjectivité du clinicien vient en interaction avec celle du patient à travers la parole. Celle-ci ne "signifie" pas de manière univoque. Les mots ne renvoient pas à des choses mais à d'autres mots qui ont des échos différents pour chacun des deux interlocuteurs selon les corrélations qui s'effectuent chez eux de manière inconsciente.

Si l'on considère la psychiatrie comme une science humaine, il est possible, voire nécessaire, de trouver dans le domaine de la psychiatrie une **place fondamentale pour la subjectivité**. La subjectivité serait même au centre du domaine psychiatrique à tel point qu'on a dit que la psychiatrie est "la science du sujet", à ceci près que les termes sont contradictoires puisqu'il n'y a de science que du général et que la subjectivité est par nature particulière.

Si la psychiatrie est centrée sur le sujet, elle ne peut exclusivement relever de la science qui suppose l'étude de phénomènes reproductibles. Or l'intersubjectivité guide la rencontre soignant-soigné et le diagnostic s'élabore et peut évoluer en même temps que s'initie et s'élabore le processus thérapeutique.

Pour J.S Strauss, notre regard reste centré sur la maladie, alors que le diagnostic et le pronostic relèvent autant du sujet souffrant que du regard du psychiatre et de leur subjectivité réciproque. Ainsi, au lieu de nier la subjectivité, il conviendrait d'aider le psychiatre à mieux cerner son attitude face à la maladie et au malade [27].

La subjectivité présente un grand intérêt, tant au niveau des évaluations, que dans une démarche générale de recherche [28]. La vérité scientifique est insuffisante en psychiatrie et il faut également envisager la vérité subjective.

C'est pourquoi, depuis les années 1990, une démarche de recherche, plus qualitative, est valorisée par certains méthodologistes d'Amérique du nord [29,30].

Et l'on voit, parmi la floraison des larges études épidémiologiques quantitatives, émerger aux États-Unis un retour vers l'analyse qualitative du cas unique. Témoigne aussi de cet intérêt pour la subjectivité, l'utilisation accrue dans les essais thérapeutiques d'échelles comme les ICG (impressions cliniques globales de R. Von Frenckell) [31], développées aux États-Unis, qui reposent sur l'évaluation subjective globale de l'état du patient par le clinicien, et comptent parmi les instruments les plus sensibles au changement. Les ICG peuvent s'utiliser dans les recherches concernant tous les psychotropes, et ce, aussi bien chez l'enfant que chez l'adulte.

Parce que la subjectivité s'inscrit dans une dynamique relationnelle, elle est source de richesse en permettant, si elle est prise en compte et analysée, de mieux comprendre le sens des comportements observés, des symptômes [28].

C'est le clinicien qui, à partir du travail réalisé avec le patient et sa famille, va aider à dégager le sens des symptômes et comportements cotés par les échelles, sens qui doit faire l'objet d'une discussion.

Le point de vue interactif sous-tendu par l'approche psychanalytique, amène les chercheurs à prendre en compte des facteurs difficilement quantifiables, voire non mesurables car ils appartiennent à la subjectivité de l'observateur et encore plus du clinicien. Une démarche plus qualitative que quantitative prend ainsi progressivement son importance.

Ainsi, J. Rosenfeld [32] montre l'intérêt de prendre en compte les éprouvés affectifs, avec sentiment de bizarrerie, des observateurs face à des films familiaux (analysés en insu) avec des enfants autistes. De Roten [33], à travers une recherche dont la méthodologie s'inspire des principes de l'EBM, reconnaît l'intérêt de l'évaluation subjective des observateurs : ce sont eux qui pointent la grande insatisfaction des parents de la population clinique, alors que ces parents eux-mêmes affirment un même degré de satisfaction que la population de contrôle.

En psychiatrie, l'instrument principal de recueil des données de recherche est le chercheur lui-même, en tant qu'être humain, seul instrument capable à ce jour de s'adapter à la situation telle qu'elle se déroule et aux imprévus, d'alterner analyses inductives et déductives, et d'utiliser des

méthodes systématiques mais également un processus d'attention approfondie et de compréhension empathique [34].

Contrairement à la médecine somatique, la médecine de la souffrance psychique doit faire appel dans sa démarche sémiologique à l'intersubjectivité qui devient moyen de connaissance. Plus encore, comme le suggère C. Deloro, ne pas envisager la subjectivité, reviendrait à détruire l'objet étudié [35].

Pour conclure, on peut donc estimer que, dans la démarche de recherche scientifique, la subjectivité n'est donc pas seulement une difficulté en terme de méthodologie, mais également un outil nécessaire pour répondre à la question incontournable du sens des symptômes en psychiatrie.

Il apparaît donc nécessaire dans les études s'intéressant au psychisme humain d'utiliser la subjectivité. Afin d'assurer au mieux la rigueur et la reproductibilité des résultats indispensables à la méthode scientifique, la mesure de la subjectivité au cours de ces études apparaît nécessaire.

III. Les échelles et leur validation

Après avoir expliqué l'origine de l'utilisation des échelles en pédopsychiatrie, nous rappellerons les principales qualités qu'elles doivent comporter pour être validées.

A. Définition et intérêt des échelles en pédopsychiatrie

L'évaluation quantitative en psychiatrie n'est pas un phénomène récent chez l'enfant. Au début du XX^e siècle, Alfred Binet, psychologue, et Théodore Simon, médecin, ont tenté de cerner et de mesurer certaines aptitudes intellectuelles de l'enfant [36]. La naissance des tests psychométriques a ouvert la voie à la notion de mesure en psychopathologie. Depuis trente ans ce courant s'est très largement développé en psychiatrie de l'adulte. La nécessité de soumettre l'évaluation clinique aux exigences des méthodologies nouvelles, notamment dans l'étude de l'effet des médicaments

psychotropes, a suscité l'intérêt pour le développement d'instruments de mesure quantifiés. Mais, alors même que la quantification est ancienne chez l'enfant, bien peu d'instruments sont actuellement disponibles en langue française pour cette période de la vie. La publication en 1985 d'un numéro spécial de la revue *Psychopharmacology Bulletin*, consacré aux échelles d'évaluation chez l'enfant, révélait pourtant un nombre non négligeable d'instruments publiés en langue anglaise. Plus tard, la revue *European Child and Adolescent Psychiatry* (1992), et la publication d'un ouvrage de Barthélémy et Lelord (1991) [37] sur ce sujet ont confirmé l'intérêt croissant pour les échelles d'évaluation en psychiatrie de l'enfant.

Le caractère limité de l'emploi de ces échelles en France tient tout autant aux difficultés méthodologiques de leur traduction et de leur validation qu'à la réticence de nombreux psychiatres d'enfant [38]. Les échelles d'évaluation ont été et restent considérées comme une approche réductrice de l'entretien avec l'enfant et sa famille. Le principal reproche qui leur a été adressé est de ne pas prendre en compte la dimension relationnelle de l'entretien pour s'en tenir à une tentative d'objectivation de symptômes le plus souvent comportementaux. Ces objections ne prennent pas en considération l'objectif fondamental de la psychopathologie quantitative, qui est de pouvoir comparer les mesures entre elles d'un sujet à l'autre et de répéter ces mesures au cours de l'évolution chez un même sujet ou un groupe de sujets. L'utilisation d'un instrument ne saurait se substituer à l'entretien clinique mais elle répond à des objectifs précis concernant l'évaluation notamment, mais non exclusivement, dans des finalités de recherche.

L'évaluation en psychiatrie de l'enfant fait référence à deux systèmes différents [39] :

- les entretiens standardisés : ces instruments correspondent à un modèle **catégoriel** des troubles mentaux de l'enfant. Ils visent à recueillir par un ensemble de questions précises les informations destinées à objectiver des critères diagnostiques. Ce sont donc avant tout des instruments **diagnostiques**. On distingue les questionnaires structurés dont les questions sont formulées de façon très précise et les questionnaires semi-structurés dont les questions autorisent une certaine liberté à l'interviewer. Nous n'aborderons pas en détail les différents questionnaires standardisés, car ils n'appartiennent pas à proprement parler au domaine des échelles d'évaluation ;

- les échelles d'évaluation : ce sont des instruments **dimensionnels** qui quantifient la symptomatologie selon une dimension continue. Il ne s'agit pas d'instruments diagnostiques.

B. Caractéristiques métrologiques [40]

Les caractéristiques métrologiques d'un instrument concernent la nature même de l'évaluation et les conditions de fiabilité et de validité de cet instrument. La connaissance de ces éléments est indispensable à l'utilisation rationnelle d'un instrument. La construction d'une échelle d'évaluation est une tâche longue et difficile. Voici les principales caractéristiques qu'il faut considérer pour une échelle d'évaluation :

1. Sensibilité, spécificité

La sensibilité et la spécificité font référence à la précision de la mesure. La sensibilité d'un instrument est sa capacité à mesurer des différences entre les sujets pour une pathologie donnée. La spécificité est la propriété de mesurer les différences entre les pathologies.

2. Fiabilité, fidélité

La fiabilité est la propriété de répétition de la mesure. La fidélité test-retest correspond à la stabilité de la mesure dans le temps chez un même patient, en l'absence de changement cliniquement perceptible de son état. La fiabilité inter-juges est calculé par la comparaison de la mesure du même patient par deux cotateurs différents. Nous développerons l'étude de cette caractéristique dans un chapitre consacré.

3. Validité

La validité d'un instrument correspond à la nature même et au sens de la mesure.

La validité externe est appréciée en comparant les résultats obtenus avec l'instrument étudié à ceux obtenus avec un autre instrument déjà validé dans le type de pathologie étudiée.

La validité interne correspond à la nature homogène ou non de la mesure. Elle fait appel aux méthodes d'analyse factorielle, dont le but est de tester le caractère mono ou plurifactoriel de l'instrument. Autrement dit, l'interprétation des scores obtenus doit tenir compte de la structure factorielle de l'instrument. Ainsi, une échelle plurifactorielle doit être interprétée en fonction des scores obtenus pour chacun des facteurs et non en fonction du score global de l'échelle. La stabilité factorielle de l'instrument est la constance de la structure factorielle lorsque certains paramètres extérieurs comme les facteurs culturels, socio-économiques ou ethniques sont modifiés. Des modifications liées à ces types de facteurs peuvent limiter l'utilisation de l'instrument.

4. Consistance interne

Elle peut être aussi appelée homogénéité et correspond à une corrélation des différents items entre eux ainsi qu'avec le score total. Elle permet de vérifier que les différents items mesurent bien la même chose. Différentes techniques statistiques peuvent être utilisées, en particulier le coefficient alpha de Cronbach, ou bien l'analyse multi factorielle.

Ces différentes modalités de construction d'un instrument soulignent la difficulté et le travail qu'impose une validation contrôlée d'une échelle. Une difficulté supplémentaire surgit lors de l'adaptation d'instruments conçus et utilisés dans un autre pays, avec la prise en compte des facteurs socioculturels.

Pour l'ensemble de ces raisons, **les échelles d'évaluation n'ont pas une validité théorique absolue, mais une validation expérimentale et empirique.**

En d'autres termes, l'instrument idéal répondant à tous les critères de validité n'existe pas, ce qui implique que chaque instrument doit être utilisé dans des conditions relativement précises et ses résultats interprétés en fonction de ces situations. Enfin, la plupart de ces instruments ont fait l'objet de plusieurs révisions, ce qui démontre leur caractère évolutif et perfectible.

C. Nature et modalité de la mesure

Les échelles d'évaluation peuvent être classées selon deux caractéristiques fondamentales :

- *Nature de la mesure* : on distingue schématiquement les échelles globales d'évaluation, qui s'intéressent à la symptomatologie psychiatrique et comportementale globale sans distinction syndromique, et les échelles syndromiques, spécifiques d'une entité donnée.
- *Modalités de la mesure* : différentes modalités sont possibles en fonction de la source utilisée pour le recueil de l'information.

L'auto-évaluation fait appel à l'enfant lui-même, qui doit choisir pour chaque question ou item entre différentes propositions pour caractériser son état actuel. La formulation et le vocabulaire doivent être adaptés à la compréhension linguistique et au niveau de développement de l'enfant. Elle nécessite une explication, voire un accompagnement pour que l'on soit assuré que les consignes sont comprises et respectées. Malgré ses limitations à cet âge, l'auto-évaluation a un réel intérêt.

L'hétéro-évaluation correspond à une information recueillie, soit par le médecin ou l'équipe soignante lors d'un entretien ou au cours d'une période d'observation, soit par l'entourage de l'enfant (parents, enseignants...) dans des conditions habituelles de vie de l'enfant. **Ce type d'évaluation nécessite des qualités précises concernant la validité inter-juges.**

Chez l'enfant, l'hétéro-évaluation a un triple avantage :

- elle permet de réunir des informations provenant de différentes sources et de déterminer, à partir du jugement clinique, les données les plus pertinents ;
- elle permet d'adapter l'entretien à l'âge de l'enfant et à son niveau de développement ;
- elle évite toute référence au niveau de lecture atteint par l'enfant évalué.

Chez l'enfant, ces deux types d'instruments sont complémentaires. Ils sont utiles pour évaluer les fluctuations comportementales et symptomatiques.

IV. La fidélité inter-juges [41 à 44]

A. Définition et intérêt

Dans des situations d'évaluation ou de recherche, il est fréquent que des individus ou des objets soient classés dans un ensemble de catégories, définies sur la base d'un certain critère et qui constituent généralement les modalités d'une variable nominale ou ordinale.

Pour vérifier l'objectivité de la démarche, celle-ci peut être réalisée de manière indépendante par deux ou plusieurs "juges", considérés comme étant également aptes à effectuer l'opération. On dispose alors de deux ou plusieurs classements des mêmes éléments, dont on souhaite déterminer le degré de concordance entre les juges qui les ont produits. On utilise également les termes d'accord, d'agrément ou de fidélité inter-juges.

L'accord entre des jugements est défini comme la conformité de deux ou plusieurs informations qui se rapportent au même objet. Cette notion implique l'existence d'une liaison entre les variables, exige des variables de même nature et un appariement des jugements.

Ce problème est souvent résolu en appliquant des méthodes corrélationnelles (calcul de coefficients de corrélation), dont le choix dépend du nombre de classements dont on dispose (deux ou plusieurs) ainsi que des caractéristiques métriques des échelles utilisées pour effectuer ces classements.

Les échelles quantitatives concernent des classements de valeurs mesurées quantifiables, par opposition aux échelles qualitatives qui s'intéressent aux mesures non quantifiables qui peuvent être rangées par ordre (ordinales) ou non (nominales).

Dans le cas de deux classements on aura recours:

- au coefficient r de Bravais-Pearson, en ce qui concerne les échelles quantitatives
- au coefficient ρ de Spearman, pour les échelles ordinales
- au coefficient Kappa de Cohen, pour les échelles nominales, que nous allons étudier dans le chapitre suivant.

Si le nombre de juges est supérieur à deux, il existe alors des méthodes qui généralisent les précédentes, comme par exemple le coefficient de concordance de Kendall ou le coefficient de corrélation intra-classe. Enfin pour les échelles nominales, on utilise le coefficient Kappa de Fleiss.

B. Le coefficient Kappa de Cohen [45 à 50]

La rigueur de la méthode scientifique réside notamment dans le fait qu'elle tient compte des effets du hasard.

Ainsi, si l'on souhaite mesurer la concordance entre deux observations réalisées par deux juges différents à partir du même outil, il faut tenir compte de la concordance qui serait uniquement due au fait du hasard.

Le test non paramétrique Kappa de Cohen ou coefficient Kappa est le test statistique le plus utilisé dans les études de reproductibilité qui nécessitent d'estimer l'agrément entre deux ou plusieurs cotations lorsqu'on étudie une variable discontinue.

Il permet de réaliser une séance de concordance entre les médecins pour estimer leur taux d'accord par le coefficient Kappa et d'étudier leurs désaccords pour y remédier.

1. Définition

Il fut proposé par Cohen [45] en 1960 et décrit ainsi :

L'accord observé entre des jugements qualitatifs ou non, résulte de la somme d'une composante "aléatoire" et d'une composante d'accord "véritable".

Le coefficient Kappa propose de chiffrer l'intensité ou la qualité de l'accord réel entre des jugements qualitatifs appariés.

Il exprime une différence relative entre la proportion d'accord observée P_o et la proportion d'accord aléatoire P_e qui est la valeur espérée sous l'hypothèse nulle d'indépendance des jugements, divisée par la quantité disponible au-delà de l'accord aléatoire.

En définitive, K est un pourcentage de l'accord maximum corrigé de ce qu'il serait sous le simple effet du hasard.

La valeur vraie du coefficient Kappa dans la population est une variable aléatoire qui suit approximativement une loi de Gauss de moyenne K et de variance $\text{Var}(K)$. L'hypothèse nulle H_0 est $K = 0$ contre l'hypothèse alternative $H_1 : K > 0$.

Dans le cas d'une étude d'accord entre deux observateurs statistiquement indépendants ayant r modalités de jugement, avec $r > \text{ou} = 2$, le coefficient Kappa s'écrit :

$$K = \frac{P_o - P_e}{1 - P_e}$$

avec, P_o : la proportion d'accord observée.

P_e : la proportion d'accord aléatoire ou concordance attendue sous l'hypothèse d'indépendance des jugements.

Les tableaux I et II présentent la notation utilisée lorsque les données sont présentées dans un tableau de contingence :

Tableau I - *Proportions jointes des jugements de deux juges sur une échelle avec r catégories*

		Juge A				
Juge B	Catégorie	1	2	...	r	Total
		1	p_{11}	p_{12}	...	p_{1r}
	2	p_{21}	p_{22}	...	p_{2r}	$p_{2.}$
	.					
	.					
	r	p_{r1}	p_{r2}	...	p_{rr}	$p_{r.}$
	Total	$p_{.1}$	$p_{.2}$...	$p_{.r}$	1

Tableau II - *Effectifs joints des jugements de deux juges sur une échelle avec r catégories*

		Juge A				
Juge B	Catégorie	1	2	...	r	Total
		1	n_{11}	n_{12}	...	n_{1r}
	2	n_{21}	n_{22}	...	n_{2r}	$n_{2.}$
	.					
	.					
	r	n_{r1}	n_{r2}	...	n_{rr}	$n_{r.}$
	Total	$n_{.1}$	$n_{.2}$...	$n_{.r}$	n

On appelle concordance observée P_o , la proportion des individus classés dans les cases diagonales de concordance du tableau de contingence, soit la somme de ces **effectifs diagonaux** (les accords) divisée par la taille de l'échantillon (n).

$$P_o = \sum_{i=1}^r p_{ii} = \frac{1}{n} \sum_{i=1}^r n_{ii}$$

avec r : le nombre de modalités de jugement.

Et la concordance aléatoire P_e qui est égale à la somme des produits des **effectifs marginaux** divisée par le carré de la taille de l'échantillon.

$$P_e = \sum_{i=1}^r p_{i.} \cdot p_{.i} = \frac{1}{n^2} \sum_{i=1}^r n_{i.} \cdot n_{.i}$$

Le coefficient Kappa est un nombre réel, sans dimension, compris **entre -1 et 1**.

On distingue 3 situations possibles pour les valeurs du Kappa théorique :

- $K = 1$, qui correspond à un accord parfait entre les juges (lorsque $P_o = 1$ et $P_e = 0,5$)
- $K = 0$, qui correspond à l'indépendance des jugements (lorsque $P_o = P_e$)
- $K = -1$, qui correspond à un désaccord total entre les juges (lorsque $P_o = 0$ et $P_e = 0,5$)

2. Significativité statistique

Pour chaque calcul statistique, il existe une probabilité de se tromper. De manière générale, un résultat est dit statistiquement significatif lorsque la probabilité de se tromper est suffisamment

faible. Le seuil de significativité est estimé par des méthodes mathématiques dépendant du test utilisé.

La détermination de la significativité du coefficient Kappa de Cohen, permet de conclure sur la probabilité que les résultats obtenus soient justes.

Il s'agit de tester l'hypothèse nulle que les jugements sont indépendants, c'est à dire que la seule liaison entre les jugements est due au hasard. Pour ce faire, Fleiss, Cohen et Everitt [46], ont montré que l'erreur standard de la concordance aléatoire S_{K0} est estimée par :

$$S_{K0} = \frac{1}{(1 - P_e) \sqrt{n}} \sqrt{P_e + P_e^2 - C} \quad \text{avec,} \quad C = \sum_{i=1}^r p_{i.} p_{.i} (p_{i.} + p_{.i})$$

Cette estimation de l'erreur standard ne requiert aucune hypothèse sur les marginales et suppose seulement n fixé.

A partir de l'erreur standard on peut calculer la probabilité p , pour laquelle l'hypothèse nulle est vraie. Il s'agit d'effectuer le rapport entre le coefficient Kappa et l'erreur standard, qui suit une loi normale centrée réduite.

Pour une probabilité $p < 0,05$, on peut conclure à une significativité statistique du coefficient Kappa calculé.

3. Remarques générales

- Il est possible que le tableau de contingence soit creux, c'est à dire qu'un des juges n'ait pas relevé du tout la présence d'une certaine catégorie de jugement; alors le calcul du Kappa ou de sa variance est impossible et l'on obtient parfois que $P_o = P_e = 1$ ou encore que $S_{K0} = 0$.

Le premier cas signifie qu'il y a eu accord parfait car $P_o = P_e = 1$ n'est possible que si une seule case de la diagonale du tableau de contingence est pleine.

Le second cas est dû au fait que la formule de la variance du Kappa ne supporte pas certaines configurations de tableaux creux.

De manière plus générale, le coefficient Kappa réagit mal au cas où l'une des modalités est rare dans la population observée.

- Fleiss [46], conseille que la taille de l'échantillon de l'étude (n) soit supérieure ou égale à $3r^2$ pour comparer deux coefficients Kappa observés et $n > 16r^2$ pour déterminer l'intervalle de confiance du Kappa.

4. Interprétation de la valeur du Kappa

Habituellement, on utilise le barème proposé par Landis et Koch [42] en 1977, mais cette grille de lecture est très contestée [51].

Cela est en bonne partie dû au fait qu'elle ne prend pas compte le fait que l'interprétation du Kappa de Cohen doit être modulée par le nombre de stades possibles.

En effet, le Kappa aura toujours tendance à être plus faible pour un grand nombre de stades possibles, que pour seulement deux stades à attribuer [52]: il est plus facile de se mettre d'accord sur "Plutôt Noir / Plutôt Blanc" que sur toute une palette comportant 5000 gris !

Ainsi, par exemple, un Kappa = 0.40 pourra être considéré comme très médiocre si deux juges avaient seulement à choisir entre deux scores A et B, mais pourra être perçu comme relativement honorable s'ils devaient choisir entre 10 stades différents.

La table de Landis et Koch [42] ne fournit donc qu'un guide de lecture, **à compléter par l'expertise du praticien** :

Tableau III - *Degré d'accord et valeur de Kappa, dans la version originale en anglais (1); accompagné de deux différentes traductions françaises retrouvées dans des études récentes [53, 54]* :

Kappa	Accord
0,80 - 1,00	Almost perfect
0,60 - 0,80	Substantial
0,40 - 0,60	Moderate
0,20 - 0,40	Fair
0,00 - 0,20	Slight
< 0,00	Poor

Traduction 1	Traduction 2
Presque parfait	Très bon
Fort	Bon
Modéré	Moyen
Faible	Médiocre
Très faible	Mauvais
Désaccord	Exécrable

Remarques :

- Ces différences de traduction en langue française, mettent en relief l'existence d'une variabilité dans l'interprétation de la valeur du Kappa.
- Le Kappa de Cohen n'est pas à proprement parler un test et ne fournit aucun aspect décisionnel. Tout comme un coefficient de corrélation entre deux variables quantitatives, il ne fournit qu'un indicateur numérique de ce que l'on cherche à mesurer, et son interprétation reste subjective.
- Pour considérer l'ensemble des catégories de jugement, on détermine la valeur du coefficient Kappa global, en calculant la moyenne des coefficients Kappa catégoriels.
- Pour chaque catégorie de jugement, il est possible de définir la valeur du Kappa par conséquent dit "catégoriel". Le calcul des Kappa catégoriels permet une étude plus fine des éventuelles discordances entre les juges.

5. Utilisation pratique

a) Outils en ligne

Pour effectuer le calcul du coefficient Kappa de Cohen, nous proposons quelques outils gratuits disponibles en ligne (ou à télécharger):

- une feuille Excel prenant en argument le tableau de contingence des scores attribués par les deux juges : <http://www.er.uqam.ca/nobel/r30574/Calcul/Kappa.xlsx>

- une plateforme de calcul en ligne prenant également en argument un tableau de contingence : <http://faculty.vassar.edu/lowry/kappa.html>

- un logiciel simple et documenté pour le calcul du Kappa fonctionnant sous DOS : http://kappa.chez-alice.fr/Kappa_cohen.htm

b) Logiciels de Statistiques

Le logiciel MedCalc peut être utilisé pour le calcul du Kappa, des explications sont disponibles à cette adresse : <http://www.medcalc.org/manual/kappa.php>

Le logiciel R dispose de nombreux packages incluant une fonction pour calculer le Kappa de Cohen, notamment :

- la fonction `ckappa` du package `psy`, qui prend en argument un tableau à deux colonnes contenant les données brutes concernant les observations des juges (donc, la liste des stades attribués par chacun d'entre eux) ;
- la fonction `kappa2` du package `irr`, qui fonctionne de la même manière que la précédente ;
- la fonction `iaa.kappa` du package `UCS`, qui prend en argument un tableau de contingence.

Le logiciel AgreeStat ne requiert aucune installation. Il s'agit d'un programme simple d'emploi, utilisable directement sous Excel, qui permet le calcul de nombreux coefficients de concordance dont le Kappa de Cohen. Il est disponible sur le site <http://www.agreestat.com/> en téléchargement payant.

2^{EME} PARTIE : ETUDE

Nous nous sommes intéressés à l'examen de la méthodologie de recherche utilisée dans une étude explorant la clinique des phases précoces de la schizophrénie, en investiguant plus particulièrement la fidélité inter-juges des mesures réalisées.

Cette étude, a été réalisée en 2010 par le Dr Claire Obacz au sein du service de pédopsychiatrie du Pr Bernard Kabuth à l'Hôpital d'Enfants de Nancy-Brabois. Elle constituait son travail de thèse de doctorat en médecine [55].

Lors de cette étude, les caractéristiques sémiologiques existant pendant l'enfance ou l'adolescence ont été recensées à partir des dossiers de 50 patients schizophrènes adultes.

Leur regroupement selon une approche dimensionnelle, et après lecture des données actuelles de la littérature, a permis de constituer une grille de lecture des dossiers psychiatriques explorant la clinique des phases précoces de la schizophrénie.

Cette grille de lecture constitue donc un instrument de mesure complexe (IMEC) des signes prémorbides et prodromiques de la schizophrénie, selon une approche dimensionnelle.

Le lien de causalité entre les signes retrouvés et la pathologie schizophrénique n'est pas établi et nécessiterait d'être testé par d'autres études longues et coûteuses.

Cependant cette exploration des dossiers de patients schizophrènes, réalisée dans l'étude du Dr Obacz, constitue un modeste pas dans la recherche psychiatrique et plus précisément dans l'exploration des signes précurseurs de la schizophrénie puisqu'il étudie les divers signes psychiques plus ou moins précoces présentés par les patients souffrant de cette pathologie.

Dans un souci de rigueur scientifique, propre à cette démarche de recherche, et avant de poursuivre et d'investiguer la spécificité des observations réalisées, s'est alors posée la question de la reproductibilité des observations effectuées dans cette étude et de son corollaire, la fidélité inter-juges de l'IMEC utilisé.

Après avoir rappelé brièvement l'étude initiale à laquelle nous nous intéressons dans ce travail, nous effectuerons une analyse globale de la méthodologie de recherche employée, puis nous explorerons plus spécifiquement la fidélité inter-juges de cet instrument qui est actuellement utilisé dans un cadre de recherche. Enfin, nous discuterons les résultats obtenus afin de proposer des perspectives de modification de cet instrument.

I. Etude initiale faisant l'objet de l'analyse méthodologique

A. Description de l'étude, matériel et méthode

1. Présentation de l'étude

Cette étude s'intéressait à des adultes schizophrènes, âgés de 18 à 30 ans, sélectionnés parmi la population de la file active de l'année 2008 au Centre Psychothérapique de Nancy-Laxou (C.P.N), et qui avaient bénéficié d'un suivi pédopsychiatrique avant d'avoir atteint leur 18^{ème} année.

a) Objectifs

L'objectif principal de cette étude portait sur la description des signes cliniques survenant lors de l'adolescence et de l'enfance du "futur" adulte schizophrène. Elle investiguait la nature de ces différents signes cliniques, leur fréquence et leur association durant la période dite "pré-schizophrénique".

L'objectif était donc de décrire les symptômes psychiatriques présents durant cette période, en vue de mettre en exergue un faisceau de symptômes plus spécifiques, jouant un rôle de sentinelle dans l'apparition d'un processus schizophrénique à l'âge adulte.

b) Type d'étude

Il s'agissait d'une étude par observation de nature descriptive et rétrospective. Il ne s'agissait donc pas d'une étude analytique mais d'une étude de cas.

2. Population étudiée

Les critères d'inclusion utilisés étaient :

- le suivi hospitalier et/ou ambulatoire au C.P.N lors de l'année 2008
- l'âge, en 2008, borné entre 18 et 30 ans ;
- le diagnostic de schizophrénie (CIM-10) saisi dans les dossiers informatiques adultes recueillis par le Département d'Information Médicale ;
- le suivi dans un des secteurs de psychiatrie infanto-juvénile de la région nancéenne avant l'âge de 18 ans.

Parmi les 54 patients répondant à ses critères, 50 ont été inclus dans l'étude et 4 dossiers n'ont pas été retrouvés.

Le tableau suivant décrit la répartition diagnostique des 50 patients selon la CIM-10 :

Diagnostic	F20.0	F20.1	F20.2	F20.3	F20.5	F20.6	F20.8	Total
Nombre	29	9	1	5	3	2	1	50
Pourcentage	58%	18%	2%	10%	6%	4%	2%	100%

F20.0 : schizophrénie paranoïde

F20.1 : schizophrénie hébéphrénique

F20.2 : schizophrénie catatonique

F20.3 : schizophrénie indifférenciée

F20.5 : schizophrénie résiduelle

F20.6 : schizophrénie simple

F20.8 : autres formes de schizophrénie

3. Données étudiées

Cette étude investiguait diverses données: administratives et sociales, les antécédents psychiatriques, des éléments concernant la prise en charge pédopsychiatrique antérieure, et son axe principal consistait dans l'observation de données cliniques avant le diagnostic de schizophrénie.

Ainsi, les caractéristiques sémiologiques présentées par les patients de la population étudiée ont été

recensées pendant leur enfance et leur adolescence.

Ces informations étaient retrouvées dans les descriptions cliniques effectuées par les différents intervenants et collectées dans les dossiers informatisés et dans les dossiers « papier » de pédopsychiatrie et de psychiatrie adulte.

Il s'agissait essentiellement d'observations de médecins et de psychologues. Elles étaient complétées par celles des orthophonistes, éducateurs, psychomotriciens et assistants sociaux.

Dans les dossiers des séjours hospitaliers, les observations infirmières permettaient également d'étoffer les comptes rendus médicaux.

4. Méthode de recueil

Le relevé des symptômes était effectué selon deux étapes. Dans un premier temps, les différents symptômes psychiatriques étaient notés, sans classification notable, à l'examen de chaque dossier. Par la suite, l'observation d'une répétition de symptômes similaires a incité la chercheuse à établir une grille de cotation. Cette grille répertoriait tous les symptômes psychiatriques retrouvés qui ont été regroupés selon différentes dimensions cliniques.

Deux périodes ont été distinguées dans l'histoire des patients avant apparition de la schizophrénie : la période de l'enfance définie de 0 à 11 ans inclus, et la période de l'adolescence comprise entre 12 et 18 ans inclus. Les dimensions retenues étaient les suivantes :

- Dimension comportementale et impulsivité
- Dimension d'anxiété
- Dimension physique
- Dimension cognitive
- Dimension négative
- Dimension positive /de discordance
- Dimension thymique (uniquement pendant l'adolescence)

Le contenu des dimensions utilisées, c'est à dire les items constituant la grille de cotation diffèrent selon la période envisagée afin de s'adapter au développement du sujet et de préciser aux mieux la sémiologie propre à l'enfance et à l'adolescence.

B. Analyse de la méthodologie de l'étude initiale

Nous ne reviendrons pas sur l'intérêt de cette étude sur la symptomatologie des phases précoces de la schizophrénie, mais nous intéresserons ici à examiner la méthodologie employée.

1. Démarche de recherche

La méthode de recherche qualitative est utilisée pour collecter et analyser des données qui ne peuvent être représentées par des nombres [56].

Une démarche qualitative plus que quantitative prend un intérêt tout particulier en psychiatrie, dans le cadre d'une approche de type holistique c'est à dire qui vise la compréhension d'un tout plus grand que la somme de ses parties.

Il s'agit d'une phase nécessaire à la recherche s'intéressant aux phénomènes de l'esprit car si la démarche quantitative est la seule qui permet d'établir des liens de causalité, la démarche qualitative est utile à créer une base de connaissance qui peut mener à des pistes de départ pour un programme de recherche de type quantitatif [57].

Dans la recherche qualitative, il existe essentiellement trois procédures d'investigation :

- l'observation, qui peut être plus ou moins structurée. Le chercheur définit a priori, ce qu'il souhaite observer en fonction des objectifs de la recherche et des outils qu'il souhaite utiliser : caméra, papier/crayon, magnétophone, ...
- l'entretien, où le chercheur peut proposer un thème général et ensuite par simple relance va recueillir le point de vue de son interlocuteur, ou bien lui administrer un hétéro-questionnaire.
- l'étude de documents écrits, qui consiste principalement en l'analyse de textes comme par exemple des notes médicales, les correspondances, les journaux intimes, la presse, mais peut également être adaptée à l'analyse d'images.

Cette dernière procédure est souvent **considérée comme idéale**, étant donné qu'il s'agit d'un recueil d'informations a posteriori sans intervention préalable du chercheur et **permet de diminuer le biais de reconstruction**.

Il s'agit de celle qui a été utilisée dans l'étude initiale à laquelle nous nous intéressons. En effet, le

recueil rétrospectif des données sémiologiques des phases précoces de la schizophrénie réalisé, ne concerne que des signes qui étaient déjà observés chez les patients de la population concernée, par des thérapeutes et d'autres intervenants du soin. Le biais de reconstruction est ainsi limité en minimisant l'intervention préalable du chercheur dans l'observation.

Elle permet également **d'éviter un biais majeur des études rétrospectives, le biais de remémoration** qui intervient lorsqu'on demande à des patients ou à leur famille de retrouver dans leurs souvenirs des signes présentés il y a bien des années.

Ainsi dans l'étude initiale, ce biais est évité car l'investigation concerne des éléments cliniques qui étaient inscrits de manière pérenne dans les dossiers par des professionnels de santé.

2. Type d'étude

Comme nous l'avons vu précédemment, chaque type d'étude possède ses avantages et ses inconvénients. **Le choix du type d'étude** est une étape essentielle de la recherche, car il **conditionne la force des résultats obtenus**.

La meilleure vérification des hypothèses est réalisée par les études expérimentales dans lesquelles les différents facteurs, sauf celui en considération, peuvent être maîtrisés.

Dans le domaine des maladies humaines, cette démarche pose un problème d'acceptabilité sur le plan éthique.

En psychiatrie, viennent également s'ajouter des difficultés d'ordre pratique puisque les différents facteurs influençant une maladie ne sont que peu souvent maîtrisables.

Par conséquent, **les études expérimentales ne sont pas réalisables** et il convient alors de procéder plutôt à des expériences dites « naturelles » ou des observations soigneusement planifiées (études cas-témoins, études de cohortes) comportant suffisamment d'informations sur les facteurs « parasites » pour permettre de corriger leurs effets lors des inférences. La méthode par observation paraît donc la stratégie accessible en premier pour déterminer une association.

Ces études analytiques par observation peuvent être prospectives (étude de cohorte) ou

rétrospectives (étude cas- témoins). Dans ces méthodes, on compare des groupes d'individus pour relever les différences en matière d'exposition ou de résultat.

Le choix de la stratégie analytique prospective apparaît difficilement accessible en terme de moyens financiers et humains pour une étude portant sur les phases précoces de la schizophrénie. En effet, selon les données de la Haute Autorité de Santé, la prévalence de la schizophrénie est de 1% en population générale et le début des troubles de la schizophrénie se situerait entre 15 et 25 ans [58].

Afin de réaliser une étude de cohorte; si l'on souhaite observer des différences statistiquement significatives, il faut pouvoir observer suffisamment de cas.

Pour investiguer les phases précoces de la schizophrénie, c'est à dire les symptômes prémorbides et les prodromes, il faut étudier la maladie et son apparition depuis la naissance des individus jusqu'à l'apparition des troubles caractérisés de schizophrénie.

La faible incidence de la schizophrénie rend donc difficile les études prospectives [59]. Elles nécessitent des échantillons de très grande taille et de très longues périodes d'observation, en général supérieures à 10 ans pour observer l'émergence de symptômes précoces de la schizophrénie [60]. Le suivi de 10 000 personnes pendant 40 ans serait nécessaire pour observer l'émergence de 90 cas de schizophrénie [61]. De plus, leur coût est très élevé et leur faisabilité en matière de détection précoce est critiquable [60,62].

Les études rétrospectives décrivant la prévalence et la durée des prodromes constituent donc une alternative reconnue dans la littérature scientifique et jugée comme étant adaptée pour l'étude des manifestations précoces de la schizophrénie [62].

Afin d'être réalisée, une étude de recherche doit s'adapter aux ressources dont elle dispose.

Le choix du caractère rétrospectif de l'étude permet de diminuer les coûts et d'obtenir des résultats de manière plus rapide, en utilisant une durée d'observation longue déjà réalisée.

La stratégie rétrospective par observation apparaît la plus adaptée, car elle est la première réalisable pour déterminer une association.

Par ailleurs, cette méthode permet d'exploiter la chance que représente la mise à disposition des dossiers pédopsychiatriques en lien avec les dossiers de psychiatrie adulte au sein d'un même

système informatisé. En effet, ce regroupement au C.P.N sur tout le secteur de la Meurthe-et-Moselle Sud constitue un cas unique en Lorraine.

Le choix de la méthode rétrospective constitue un biais mais constitue également la force de cette étude.

3. Données étudiées et méthodologie de recueil

Certaines remarques méritent d'être faites concernant le recueil des données dans l'étude initiale :

- Le regroupement des signes selon des dimensions cliniques peut paraître "arbitraire", car il n'existe pas aujourd'hui de consensus sur le nombre de dimensions cliniques et leurs contenus. Cependant les dimensions utilisées résultent d'une synthèse de la littérature scientifique actuelle sur les phases précoces de la schizophrénie et recouvrent un éventail vaste des divers symptômes existant en pédopsychiatrie.

- Dans l'utilisation de la grille de cotation, l'appartenance de certains items à une dimension spécifique n'est pas univoque. Notons par exemple l'item « alogie » défini par « discours pauvre avec barrages » qui pourrait convenir aussi bien à la dimension négative qu'à la dimension cognitive.

- La distinction de deux périodes uniques (enfance et adolescence) peut être critiquable. En effet, les différentes phases du développement psychologique utilisés en pédopsychiatrie font plutôt référence à la théorie psychanalytique de Freud et distinguent les périodes suivantes : 1-3 ans (stade anal), 3-6 ans (stade phallique), 6-12 ans (période de latence) et 12-18 ans (stade génital). Ces différentes périodes auraient pu être distinguées dans l'étude initiale.

II. Examen de la reproductibilité des résultats au travers de la fidélité inter juges d'un instrument de mesure complexe (IMEC)

A. Justification de l'examen

Les résultats obtenus dans l'étude initiale peuvent être soumis à caution sur leur fondement scientifique.

Dans cette étude, tous les sujets sont des patients souffrant de schizophrénie à l'âge adulte ; ce qui fait envisager un possible biais de suspicion des résultats. L'observation des éléments cliniques possiblement attribuables à la schizophrénie n'était pas effectuée en aveugle de la présence de cette maladie. Par ailleurs, la présence ou l'absence d'un signe clinique n'est pas toujours noté comme telle ; et les données sémiologiques inscrites dans les dossiers sont soumises à l'interprétation du lecteur.

Il est donc probable que l'on puisse reprocher le caractère subjectif des résultats.

Néanmoins, il s'agit d'une caractéristique propre au domaine (pédo)psychiatrique qui investigate des effets difficilement objectivables autrement que par la subjectivité.

Il est donc apparu nécessaire, d'investiguer la reproductibilité des résultats obtenus dans cette étude de recherche.

Dans un but de rigueur scientifique, nous avons donc tenté de mesurer cette part de variabilité, liée à la subjectivité.

Il a été décidé d'effectuer à nouveau le recensement des symptômes de la même population à l'aide de la grille de cotation dimensionnelle élaborée dans l'étude initiale mais cette fois ci par un autre juge.

B. Elaboration d'un instrument de mesure complexe (IMEC)

1. Justification

Les diverses informations séméiologiques recueillies dans l'étude initiale étaient réunies selon une lecture dimensionnelle de la clinique de la schizophrénie, qui a été formalisée sous la forme d'une grille de cotation.

Cette grille de cotation constituait donc les prémisses d'un outil pour effectuer un relevé systématique de symptômes psychiatriques possiblement attribuables aux phases prémorbides et prodromiques de la schizophrénie.

Afin de pouvoir permettre son utilisation avec fiabilité, il est apparu nécessaire d'envisager sa formalisation complète.

Cet outil, qui prend l'apparence d'une échelle, ne peut pour le moment, en obtenir le titre, n'étant pas validée comme telle. Nous parlerons donc d'**instrument de mesure complexe (IMEC)**.

Au travers de l'examen de la fidélité inter-juges de l'IMEC, il s'agit donc d'étudier la reproductibilité des mesures effectuées et parallèlement d'envisager une étape vers l'amélioration et la validation de cet outil de mesure.

2. Méthodologie de construction

L'évaluation quantitative en psychiatrie s'effectue à partir d'échelles préalablement validées.

Cependant, il n'existe pas à l'heure actuelle d'échelle validée en français relevant les symptômes présents lors des phases précoces de la schizophrénie.

La construction d'un instrument de mesure de ces symptômes utilisable en langue française apparaît par conséquent légitime.

Chaque instrument est unique dans sa structure et son contenu. L'utilisation d'une échelle d'évaluation ou d'un IMEC nécessite une formation préalable. Bien qu'il soit destiné à des

utilisateurs censés partager une base de connaissances commune, il fait appel à un certains nombre de présupposés qui sont habituellement décrits dans un manuel du mode de passation de l'outil. La réalisation d'un manuel est donc une étape nécessaire à la validation de cet outil. Il permet d'obtenir une standardisation, c'est à dire une uniformité de passation et de correction.

Ce manuel a été produit selon les critères du recueil des signes effectué dans l'étude initiale ; il vise à expliciter l'utilisation de l'instrument de mesure et définit les différents items des dimensions recherchées par l'IMEC.

L'IMEC a été construit selon les critères habituels des échelles d'évaluation [40]:

- présentation structurée autour de différents axes (les dimensions) ;
- aspect de « check-list » relevant la présence ou l'absence d'items distincts (les signes cliniques) ;
- détails sur l'utilisation de l'outil donnés dans un paragraphe séparé ;
- unité de l'instrument définie autour d'une même entité (la schizophrénie).

Il n'existe cependant pas de score attribuable et le nombre d'items dans chaque dimension n'est pas homogène.

3. Séances de mise au point

L'utilisateur de l'IMEC doit être familiarisé avec les descriptions des items et les critères de cotation.

Ainsi, avant le nouveau recueil de symptômes visant à évaluer la variabilité des mesures, il a été procédé à la cotation de 3 dossiers tests ne faisant pas partie des dossiers de la population étudiée. Ces dossiers ont été choisis dans la population de patients de la file active de 2009 et partageant les mêmes critères d'inclusion que ceux requis pour l'étude initiale.

Ces séances ont permis de préciser les modalités du relevé des signes et ont fait apparaître la variabilité individuelle dans l'interprétation des symptômes entre les deux juges.

On peut cependant regretter qu'il n'ait pas été défini complètement avant le relevé par le premier juge. Mais cela s'explique aisément dans cette démarche de recherche : le premier juge est également celui qui a participé à l'élaboration de l'outil et a donc utilisé une version non formalisée ; le deuxième juge a bénéficié d'un outil dans sa version finale.

4. Aspect de l'IMEC

L'instrument de mesure complexe utilisé est présenté en annexe à la page 106.

Il s'agit d'une « check-list » de 48 items constituée de différents symptômes psychiatriques retrouvés dans les dossiers. Deux parties sont distinguées correspondant à deux périodes d'investigation des phases précoces et prodromiques de la schizophrénie: l'enfance et l'adolescence.

Pour l'enfance, le nombre d'items différents utilisés est de 23, tandis que pour l'adolescence il est de 25. Ces items sont regroupés selon 7 dimensions cliniques :

- Dimension comportementale et impulsivité
- Dimension d'anxiété
- Dimension physique
- Dimension cognitive
- Dimension négative
- Dimension positive /de discordance
- Dimension thymique

Deux modalités de cotation sont proposées pour chaque item de la liste : « absent » ou « présent ».

Un manuel de passation a été formalisé et constitue le paragraphe suivant.

5. Manuel d'utilisation de l'IMEC

Les signes cotés correspondent à des symptômes relevés dans les dossiers par les différents thérapeutes: médecin, infirmier, psychologue, psychomotricien, orthophoniste... Ce sont les symptômes objectivés durant les consultations ou des symptômes rapportés par les parents et mentionnés dans les dossiers par les cliniciens.

Les manifestations peuvent être décrites dans les dossiers sans correspondre aux dénominations précises des items. L'IMEC fait appel au sens clinique de l'utilisateur. Il relève d'une certaine subjectivité propre à chacun, dont nous tenterons d'évaluer l'effet sur la mesure.

Les signes à relever sont décrits dans les dossiers pédopsychiatriques au moment où ils sont objectivés. Ils peuvent être également décrits a posteriori dans les dossiers de pédopsychiatrie ou de psychiatrie adulte et sont retenus s'ils font **explicitement** référence à un âge donné ou à une période précise. Deux périodes sont distinguées : l'enfance correspondant aux âges entre 0 et 11 ans inclus; l'adolescence correspondant aux âges entre 12 et 18 ans inclus.

Chaque item est coté « présent » s'il répond à la description faite dans le manuel de passation. Les items non retrouvés dans les dossiers ou ne répondant pas à la description sont cotés 'absents'.

Les items sont cotés pour chaque période même s'ils correspondent à des âges différents (ex : si des signes d'anxiété sont retrouvés à 4 ans, des éléments phobiques à 7 ans et des éléments obsessionnels à 15 ans sont retrouvés, alors les 2 items « anxiété » et « éléments phobiques » sont cotés présents pour l'enfance et l'item « éléments obsessionnels » est coté présent pour l'adolescence).

Aucun critère de durée des symptômes ne doit être pris en compte pour coter, seule la présence ou l'absence à un moment donné est considérée.

a) Cotations pour l'enfance : de 0 à 12 ans

Les signes comptabilisés sont ceux que le patient a présentés de la naissance jusqu'à 11 ans inclus. A partir de 12 ans, ils doivent être reportés dans la partie de l'IMEC relative à l'adolescence.

Pour chaque dimension, il s'agit de coter les items « présents » s'ils répondent à leur description présentée ci-dessous ; sinon ils sont cotés « absents ».

▪ **Dimension comportementale, d'impulsivité :**

Passages à l'acte auto ou hétéro agressifs (comprenant également les crises clastiques).

Troubles du comportement non précisés : la mention trouble du comportement apparaît dans les dossiers mais il n'y a pas de plus amples précisions.

Une catégorie « **autres** » regroupant l'impulsivité, l'intolérance à la frustration, l'opposition, l'irritabilité, l'agitation et les colères fréquentes.

▪ **Dimension d'anxiété :**

Angoisse /anxiété de séparation.

Éléments du registre obsessionnel : rituels, obsessions idéatives, ...

Éléments phobiques : phobie scolaire, phobie spécifique (agoraphobie, hydrophobie, peur du bruit, du vélo, du vide, ...). Si la phobie a un caractère étrange, auquel cas et à la lecture du dossier un item de discordance (bizarrerie) peut être rajouté.

▪ **Dimension physique :**

Anorexie/perte de poids non organique.

Troubles du sommeil : sommeil perturbé, endormissement tardif, réveils précoces, cauchemars importants, inversion du rythme nyctéméral.

Manifestations psychosomatiques : céphalées, malaises, douleurs abdominales et autres symptômes physiques. Les termes « somatisations » ou « psychosomatiques » sont précisés dans les dossiers.

Énurésie primaire ou secondaire.

- **Dimension cognitive :**

QI bas: lorsque les bilans psychologiques sont retrouvés dans les dossiers et indiquent une valeur de Quotient Intellectuel <70, ou que le sujet est orienté en Institut Médico-Educatif.

Retard du développement moteur : il existe un retard de la marche ou un retard des acquisitions. L'item est également validé lorsqu'il est noté dans le dossier « retard du développement psychomoteur » sans précisions supplémentaires.

Retard de langage, de parole.

Troubles de la latéralisation : le trouble est mentionné dans les comptes rendus médicaux ou ceux des psychomotriciens. L'item est validé également lorsque des difficultés d'orientation spatio-temporelle sont retrouvées.

Une dernière catégorie appelée « **autres** » comprenant les troubles de la mémorisation, de la concentration, de l'attention et les difficultés de raisonnement, mentionnés dans les dossiers.

- **Dimension négative :**

Retrait : item validé lorsque il est mentionné un retrait ou une inhibition majeure.

Difficultés relationnelles : difficultés à entrer en relation avec le groupe de pairs, troubles de l'adaptation à l'école, le terme « difficultés relationnelles » peut être noté dans le dossier sans plus de précisions.

Stéréotypies motrices : terme mentionné comme tel dans les dossiers.

▪ **Dimension positive/de discordance :**

Cette dimension peut paraître atypique pour décrire les symptômes de l'enfance, néanmoins elle correspond généralement aux items retrouvés chez les enfants aux antécédents de trouble envahissant du développement.

Bizarries comportementales : le terme « bizarre » est retrouvé dans les dossiers pour qualifier le comportement ou l'enfant de manière générale.

Fabulations : le terme « fabulation » est précisé dans les dossiers en tant que tel.

▪ **Dimension thymique :**

Items dépressifs : correspondant aux termes « affects dépressifs », « syndrome dépressif », « dépression » retrouvés dans les dossiers.

Idées suicidaires : qui correspondent au terme « idées suicidaires », « idées noires ».

Fluctuations de l'humeur : comprend aussi ce qui se réfère à la « labilité thymique », et à « l'excitation psychomotrice » (avec note thymique) précisées dans les dossiers.

b) Cotations pour l'adolescence : de 12 à 18 ans

Les signes comptabilisés sont ceux que le patient a présentés, à partir de l'âge de 12 ans et avant que le diagnostic de schizophrénie n'ait été porté.

Si le diagnostic a été porté après l'âge de la majorité, les manifestations cliniques présentées sont relevées jusqu'à l'âge de 18 ans inclus (âge inférieur à 19 ans).

Pour chaque dimension, il s'agit de coter les items « présents » s'ils répondent à leur description présentée ci-dessous ; sinon ils sont cotés « absents ».

Pour chaque dimension, il s'agit de coter les items « présents » s'ils répondent à leur description présentée ci-dessous ; sinon ils sont cotés « absents ».

▪ **Dimension comportementale, d'impulsivité :**

Passages à l'acte auto ou hétéro-agressifs qui comprennent également les crises clastiques, les tentatives de suicide.

Comportement délictueux : il est évoqué une délinquance, des faits de vol, ...

Dernier item appelé « **autres** », qui regroupe les manifestations suivantes : agressivité sans passage à l'acte, intolérance à la frustration, opposition, agitation, fugues et les troubles du comportement non spécifiés dans les dossiers (la mention « trouble du comportement » apparaît dans le dossier sans être détaillée).

▪ **Dimension d'anxiété :**

Angoisse/attaques de paniques : manifestations décrites ou citées dans les dossiers : anxiété, angoisse paroxystique. Les descriptions somatiques (« boules dans le ventre », « palpitations »...)

sont également comptabilisées s'il est indiqué dans le dossier qu'elles se réfèrent à une étiologie anxieuse.

Éléments du registre obsessionnel : rituels, obsessions idéatives, comportement obsessionnel.

Éléments phobiques : le terme « phobie » doit être précisé dans les dossiers. Les différentes phobies sont comprises même si elles possèdent un caractère étrange.

▪ **Dimension thymique** :

Items dépressifs : correspondant aux termes « affects dépressifs », « syndrome dépressif », « dépression » retrouvés dans les dossiers.

Idées suicidaires : qui correspondent au terme « idées suicidaires », « idées noires ».

Fluctuations de l'humeur : comprend aussi ce qui se réfère à la « labilité thymique », et à « l'excitation psychomotrice » (avec note thymique) précisées dans les dossiers.

▪ **Dimension cognitive** :

Troubles de l'attention et de la concentration : cités en tant que tels ou décrits dans les dossiers comme les difficultés attentionnelles, les difficultés à se concentrer en classe, l'attention fluctuante durant les entretiens.

Diminution des capacités d'abstraction : citée dans les dossiers en tant que telle ou les termes « difficultés de symbolisation » ou « difficultés de conceptualisation » sont employés.

Lenteur idéative : des difficultés de raisonnement sont mentionnées ou le raisonnement est jugé lent.

Dernier item dénommé « **autres** » : qui regroupe la maladresse gestuelle et les troubles des fonctions instrumentales (trouble de l'écriture, du calcul, du langage...).

▪ **Dimension négative :**

Retrait : un retrait est observé, cela se réfère également aux termes « isolement », « inhibition », « renfermé ».

Froideur affective : correspondant aux termes « affects abrasés » et/ou « détachés », « pauvreté des affects », « réticence ».

Avolition : désintérêt, diminution du rendement à l'école ou dans les différentes activités, apragmatisme, apathie. C'est-à-dire les éléments qui se réfèrent à une altération du fonctionnement habituel.

Alogie : discours pauvre, barrages.

▪ **Dimension positive de discordance :**

Discordance : bizarrerie du comportement, du contact, des propos, maniérisme, rires immotivés, diffluence de la pensée, déréalisation/dépersonnalisation.

Hallucinations (auditives, visuelles, cénesthésiques): si le terme est précisé tel quel, également lorsqu'est noté dans le dossier « entend des voix », « voit des formes, couleurs, objets, ... » ou « sensations corporelles bizarres ».

Sentiment de persécution, suspicion : allant de la forte conviction au sentiment flou de persécution, de suspicion qui peut être décrit dans les dossiers (concernant l'entourage, la famille, le thérapeute, ...); les éléments interprétatifs sont également comptabilisés.

Idées délirantes : idées de référence, idées délirantes mégalomaniaques, de culpabilité, d'empoisonnement, ou « idées délirantes » inscrites dans le dossier sans plus de précision.

- **Dimension physique :**

Anorexie/perte de poids non organique.

Troubles du sommeil : sommeil perturbé, endormissement tardif, réveils précoces, cauchemars importants, inversion du rythme nyctéméral.

Incurie : défaut d'hygiène, présentation négligée.

Manifestations psychosomatiques : céphalées, malaises, douleurs abdominales et autres symptômes physiques. Les termes « somatisations » ou « psychosomatiques » sont précisés dans les dossiers.

C. Objectifs

1. Objectif principal

L'objectif premier est de mesurer la variabilité inter-juges par le calcul des coefficients Kappa de Cohen des différents items constituant l'IMEC et de leur regroupement en dimensions.

Il s'agit d'une analyse statistique concernant la fidélité inter-juges de cet outil.

Cet examen constitue une analyse de la reproductibilité du relevé de symptômes effectué, c'est à dire de la **stabilité de la mesure**.

2. Objectif secondaire

Cette analyse doit permettre de distinguer les items pour lesquels la concordance est la plus faible afin d'envisager des modifications de l'IMEC dans sa structure et dans son contenu (formulation des items).

D. Matériel et méthodes

1. Méthodologie de revue de la littérature

La recherche bibliographique a été effectuée essentiellement en ligne sur Internet à partir de quelques banques de données médicales :

- **Medline** et **Science Direct** : banques de données médicales anglo-saxonnes, où la recherche s'effectue par mots-clés ;
- **Sudoc** : catalogue collectif français du Système Universitaire de Documentation réalisé par les bibliothèques et en centres de documentation de l'enseignement supérieur et de la recherche ;
- **Service commun de documentation de la bibliothèque de l'Université Henri Poincaré de Nancy** : revues médicales accessibles en lignes, ouvrages de référence disponibles sur place ;
- **Banque de recherche Internet Google Scholar.**

Nous avons également effectué une partie de notre recherche à la **bibliothèque du Centre Hospitalier de Lorquin** ainsi qu'à la **bibliothèque du Centre Psychothérapique de Nancy**.

2. Type d'examen

Il ne s'agit pas d'une étude au niveau épidémiologique du terme, mais d'un examen de la reproductibilité des résultats d'une étude rétrospective de cas. Cet examen est centré sur l'analyse de la fidélité inter-juges de l'IMEC construit au préalable.

Nous avons tenté d'effectuer cet examen de nature statistique, de la manière la plus adaptée, en considérant la spécificité de l'objet (pédo)psychiatrique.

3. Population étudiée

Pour tester la fidélité inter juges de l'IMEC, nous avons effectué un recensement des symptômes sur la même population, en reprenant les mêmes dossiers que dans l'étude du Dr Obacz.

Les critères d'inclusion sont donc identiques à ceux retenus dans l'étude initiale.

4. Modalités de recueil des données

Nous avons obtenu par courrier, un accord des différents médecins responsables des unités d'hospitalisation et des C.M.P rattachés au C.P.N, pour rechercher et consulter les dossiers des patients de la population étudiée.

Nous avons effectué le relevé des symptômes dans des conditions les plus proches possibles de celles de l'étude initiale.

Nous avons utilisés les mêmes dossiers informatisés et tenté d'obtenir les mêmes parties des dossiers « papier ». L'exploration des dossiers a été réalisée dans les lieux où ils étaient retrouvés :

- aux archives générales du C.P.N,
- dans les unités hospitalières du C.P.N : adultes, lorsque les dossiers de pédopsychiatrie étaient regroupés avec les dossiers adultes, ou pédopsychiatriques lorsqu'ils n'étaient pas fusionnés,
- dans les centres médico-psychologiques des secteurs de psychiatrie adulte ou infanto-juvénile.

5. Données recueillies

Notre examen concerne les données séméiologiques retrouvées par l'utilisation de l'IMEC considéré.

L'IMEC a été utilisé pour chaque patient. Les données recueillies ont ensuite été retranscrites sur Excel® pour effectuer leur traitement informatique et statistique.

6. Analyse statistique

La méthodologie employée, ainsi que l'analyse statistique, ont été réalisés avec la collaboration de statisticiens :

- du service Epidémiologie et Evaluations Cliniques du Centre Hospitalier Universitaire de Nancy Brabois (consultation du Dr N. Agrinier, tests réalisés par M-L. Erpelding) ;
- et celle du Dr F. Rouyer-Ligier, Pédopsychiatre et Praticien Hospitalier disposant d'une formation en Epidémiologie clinique et en Bio-statistiques (Master en Santé Publique),

exerçant au sein du service de psychiatrie de l'enfant et de l'adolescent de l'Hôpital d'Enfants de Nancy Brabois.

Les résultats obtenus ont été exploités sur le plan statistique afin d'étudier la fidélité inter-juges à partir de la mesure du coefficient Kappa de Cohen.

Les calculs statistiques ont été effectués à l'aide du logiciel AgreeStat 2011.2 : Chance corrected agreement & Intraclass correlation Coefficients with Excel ; ainsi que du logiciel SAS 9.3.

E. Résultats

1. Répartition des symptômes selon le juge

Afin d'évaluer le comportement des juges et de repérer des différences d'interprétation des items de l'instrument de mesure, nous nous sommes d'abord intéressés à la description de la répartition des différentes caractéristiques sémiologiques retenues selon le juge.

a) Descriptif des items relevés selon le juge

Nous présentons ci-après, pour chacun des items de l'IMEC classés par dimensions, le nombre de patients (parmi les 50 de la population) où l'item correspondant a été coté « présent » dans les dossiers par chacun des deux juges.

Pour la période de l'enfance, nous avons retrouvés les effectifs présentés dans les tableaux suivants :

ENFANCE

▪ Dimension d'impulsivité :

	Juge 1	Juge 2
passages à l'acte auto/hétéro-agressif	5	12
troubles du comportement NS	5	7
autres	13	14

▪ Dimension d'anxiété :

	Juge 1	Juge 2
anxiété/angoisse de séparation	9	15
éléments phobiques	4	6
éléments obsessionnels	1	1

▪ Dimension physique :

	Juge 1	Juge 2
anorexie	4	6
énurésie	10	11
troubles du sommeil	10	6
manifestations psychosomatiques	4	2

▪ Dimension cognitive :

	Juge 1	Juge 2
QI bas	9	6
retard du développement moteur	14	8
retard de langage	16	14
autres	12	11
troubles de la latéralisation	5	8

▪ **Dimension négative :**

	Juge 1	Juge 2
retrait	12	11
difficultés relationnelles	3	6
stéréotypies	3	8

▪ **Dimension positive/de discordance :**

	Juge 1	Juge 2
bizarries	2	2
fabulations	2	5

▪ **Dimension thymique :**

	Juge 1	Juge 2
items dépressifs	0	4
idées suicidaires	0	0
fluctuations thymiques	0	1

Pour la période de l'adolescence, nous avons retrouvés les effectifs présentés dans les tableaux suivants :

ADOLESCENCE

▪ **Dimension d'impulsivité :**

	Juge 1	Juge 2
passages à l'acte auto/hétéro-agressif	21	25
comportement délictueux	10	10
autres	26	35

▪ **Dimension d'anxiété :**

	Juge 1	Juge 2
anxiété/attaque de panique	25	32
éléments obsessionnels	9	16
éléments phobiques	2	1

▪ **Dimension thymique :**

	Juge 1	Juge 2
items dépressifs	25	25
idées suicidaires	15	2
fluctuations thymiques	7	6

▪ **Dimension cognitive :**

	Juge 1	Juge 2
trouble de l'attention / concentration	10	9
difficultés d'abstraction	3	5
lenteur idéative	13	9
autres	6	6

▪ **Dimension négative :**

	Juge 1	Juge 2
retrait	34	29
froider affective	10	15
avolition	27	16
alogie	5	15

▪ **Dimension positive/de discordance :**

	Juge 1	Juge 2
discordance	20	24
hallucinations	2	15
persécution	19	20
idées délirantes	5	21

▪ **Dimension physique :**

	Juge 1	Juge 2
perte de poids/anorexie	9	18
troubles du sommeil	16	24
incurie	8	10
manifestations psychosomatiques	4	4

b) Description des données recueillies

Afin de comparer les comportements des 2 juges concernant la cotation effectuée, nous nous sommes attachés à décrire la distribution des effectifs mesurés par ceux-ci.

La moyenne des effectifs totaux de chaque item a été établie pour chaque juge, tout comme le calcul des écart-types de ces moyennes.

La comparaison des données retrouvées par chaque juge pour la même population de l'échantillon est ensuite effectuée par la réalisation d'un test de Student apparié avec un seuil de significativité p fixé à 0,05.

	Juge 1	Juge 2
Moyenne nb patients	9,87	11,58
Ecart-type	8,12	8,60
Student (<i>p</i>)	0,032	

En moyenne, le juge 2 relève donc plus de patients répondant à la description d'un item, comparativement au juge 1, et cette différence apparaît significative sur le plan statistique.

Le juge 2 relève donc significativement plus d'items présents que le juge 1.

2. Mesure de la fidélité du relevé de chaque item entre les juges

Afin d'investiguer la concordance des jugements effectués entre les deux juges, nous avons calculer les coefficients Kappa catégoriels, c'est à dire le coefficient Kappa pour chaque item de l'instrument de mesure utilisé pour les périodes de l'enfance et de l'adolescence.

a) Coefficients Kappa catégoriels

Les tableaux ci-dessous présentent les résultats pour la période de l'enfance, des coefficients Kappa catégoriels, ainsi que leur degré de significativité *p*. Le résultat est considéré comme statistiquement significatif lorsque $p < 0,05$.

ENFANCE

- **Dimension d'impulsivité :**

	Kappa	<i>p</i>
passages à l'acte auto/hétéro-agressif	0,52	0,000
troubles du comportement NS	0,43	0,002
autres	0,44	0,002

▪ **Dimension d'anxiété :**

	Kappa	<i>p</i>
anxiété/angoisse séparation	0,46	0,001
éléments phobiques	0,56	0,000
éléments obsessionnels	-0,02	0,885

▪ **Dimension physique :**

	Kappa	<i>p</i>
énurésie	0,46	0,001
anorexie	0,56	0,000
troubles du sommeil	0,12	0,384
manifestations psychosomatiques	0,65	0,000

▪ **Dimension cognitive :**

	Kappa	<i>p</i>
QI bas	0,77	0,000
retard du développement moteur	0,66	0,000
retard de langage	0,81	0,000
autres	0,72	0,000
troubles de la latéralisation	0,56	0,000

▪ **Dimension négative :**

	Kappa	<i>p</i>
retrait	0,72	0,000
difficultés relationnelles	0,15	0,240
stéréotypies	0,50	0,000

▪ **Dimension positive/de discordance :**

	Kappa	<i>p</i>
bizarries	0,48	0,001
fabulations	0,55	0,000

▪ **Dimension thymique :**

	Kappa	<i>p</i>
items dépressifs	nc	nc
idées suicidaires	nc	nc
fluctuations thymiques	nc	nc

Les tableaux ci-dessous présentent les résultats pour la période de l'adolescence, des coefficients Kappa catégoriels, ainsi que leur degré de significativité *p*. Le résultat est considéré comme statistiquement significatif lorsque $p < 0,05$.

ADOLESCENCE

▪ **Dimension d'impulsivité :**

	Kappa	<i>p</i>
passages à l'acte auto/hétéro-agressif	0,44	0,002
comportement délictueux	0,88	0,000
autres	0,31	0,019

▪ **Dimension d'anxiété :**

	Kappa	<i>p</i>
anxiété/attaque de panique	0,40	0,003
éléments obsessionnels	0,53	0,000
éléments phobiques	-0,03	0,836

▪ **Dimension thymique :**

	Kappa	<i>p</i>
items dépressifs	0,52	0,000
idées suicidaires	0,18	0,027
fluctuations thymiques	0,03	0,841

▪ **Dimension cognitive :**

	Kappa	<i>p</i>
trouble de l'attention / concentration	0,42	0,003
difficultés d'abstraction	0,46	0,001
lenteur idéative	0,42	0,002
autres	0,05	0,708

▪ **Dimension négative :**

	Kappa	<i>p</i>
retrait	0,45	0,001
froideur affective	0,53	0,000
avolition	0,42	0,001
alogie	0,18	0,123

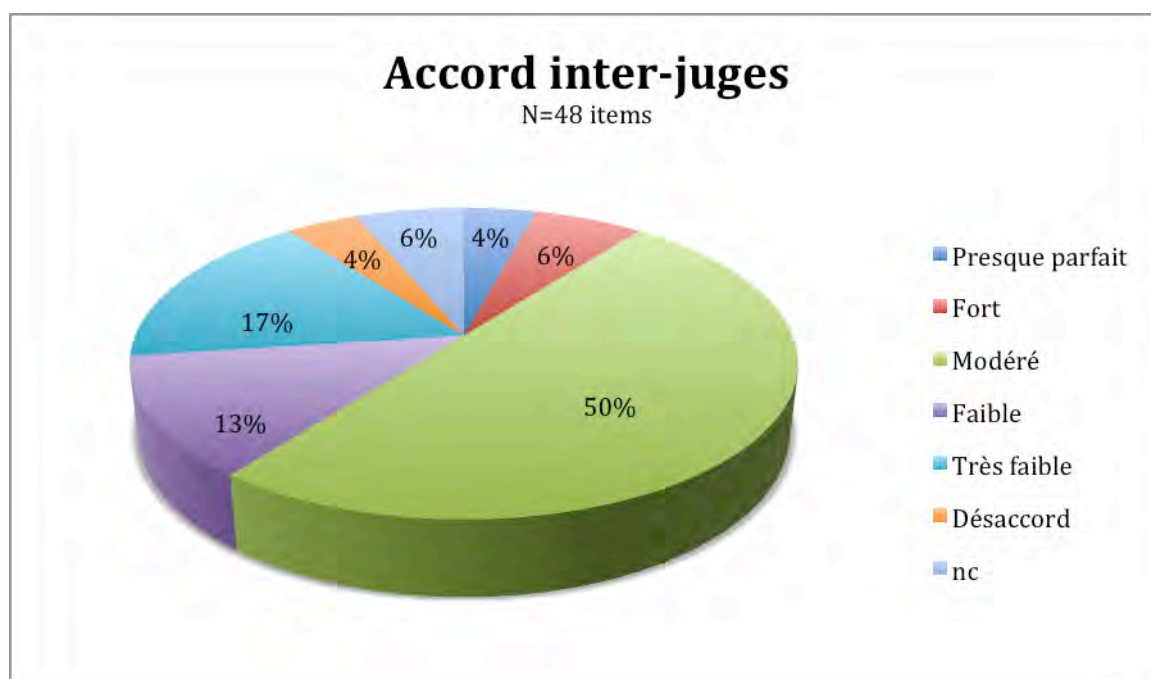
▪ **Dimension positive/de discordance :**

	Kappa	<i>p</i>
discordance	0,68	0,000
hallucinations	0,05	0,529
persécution	0,45	0,001
idées délirantes	0,17	0,070

▪ **Dimension physique :**

	Kappa	<i>p</i>
perte de poids/anorexie	0,37	0,004
troubles du sommeil	0,35	0,009
incurie	0,46	0,001
manifestations psychosomatiques	0,46	0,001

Nous avons relevé, selon la table utilisée par Landis et Koch [détaillée dans la partie théorique IV.B.4], le nombre d'items pour lesquels la valeur de Kappa indique un même degré d'accord et nous avons fait figurer en pourcentage ce que cette valeur représente sur les 48 items recensés par l'IMEC.



Plus de la moitié des items de l'IMEC sont donc observés avec un accord inter-juges considéré comme modéré (coefficient Kappa de Cohen compris entre 0,40 et 0,60). Cela est vérifié aussi bien pour la période de l'enfance (11 items) que pour celle de l'adolescence (13 items).

b) Coefficient Kappa global de l'IMEC

L'IMEC est constitué de 48 items repartis dans deux périodes distinctes. Pour 3 items, le coefficient Kappa de Cohen n'a pu être établi.

Par conséquent, afin d'obtenir l'estimation d'un coefficient Kappa global de l'IMEC, nous avons établi la moyenne des 45 coefficients Kappa catégoriels obtenus précédemment.

La valeur retrouvée est alors, Kappa global de l'IMEC = 0,43.

c) Limites

L'estimation du coefficient Kappa est soumise à une marge d'erreur ; le calcul de la significativité p permet de déterminer les mesures pour lesquelles cette marge d'erreur ne permet pas de conclure sur la fiabilité du Kappa.

On retient comme seuil $p < 0,05$ afin de retenir un coefficient Kappa comme statistiquement significatif.

Cela concerne donc un certain nombre de résultats qui doivent être considérés comme non significatifs sur le plan statistique et que nous avons identifiés en italique dans les tableaux du paragraphe 2.b.

Le calcul du Kappa pour les items de la dimension thymique à l'enfance n'a pu être effectué (nc) car le juge 1 n'a relevé aucun symptôme de cette dimension pendant cette période.

3. Mesure de la fidélité du relevé par dimensions

L'instrument de mesure s'intéresse à relever des symptômes au sein de dimensions distinctes. L'étude initiale suggérait que les dimensions étaient pérennes au cours du développement de l'individu [55]. En conséquence, nous avons étudié la stabilité des résultats dans l'existence de ce regroupement dimensionnel.

La fidélité inter-juges pourrait donc être sensiblement meilleure si on l'examine au niveau de ces dimensions plutôt qu'au niveau des items les constituant.

Chaque dimension a été considérée comme validée lorsqu'au moins un des items la constituant était retrouvée par le juge.

Un calcul du coefficient Kappa a par la suite été calculé selon la même formule statistique utilisée pour les Kappa catégoriels :

a) Coefficients Kappa dimensionnels

Le tableau ci-dessous présente les résultats pour la période de l'enfance, des coefficients Kappa dimensionnels, ainsi que le degré de significativité p fixé au seuil $p < 0,05$.

ENFANCE

Dimensions	Kappa	p
Impulsivité	0,69	0,000
Anxiété	0,65	0,000
Physique	0,55	0,000
Cognitive	0,87	0,000
Négative	0,72	0,000
Positive	0,46	0,001
Thymique	nc	nc

Le tableau ci-dessous présente les résultats pour la période de l'adolescence, des coefficients Kappa dimensionnels, ainsi que le degré de significativité p fixé au seuil $p < 0,05$.

ADOLESCENCE

Dimensions	Kappa	p
Impulsivité	0,42	0,001
Anxiété	0,58	0,000
Thymique	0,31	0,030
Cognitive	0,43	0,003
Négative	0,46	0,001
Positive	0,54	0,000
Physique	0,50	0,000

b) Analyse

Le calcul du coefficient Kappa pour la dimension thymique pendant l'enfance ne peut être calculé (nc) car le juge 1 n'a recherché aucun symptôme dans cette dimension.

Dans cette configuration de regroupement dimensionnel, l'estimation du coefficient Kappa est meilleure sur le plan statistique et tous les résultats apparaissent significatifs ($p < 0,05$). Cela signifie que l'estimation de la fidélité inter-juges apparaît plus valide selon le regroupement dimensionnel.

Hormis dans la dimension thymique, les coefficients Kappa dimensionnels retrouvés sont supérieurs à 0,40 ; ce qui signifie que **les juges ont un accord considéré au minimum comme modéré** si l'on considère le regroupement dimensionnel.

Pendant l'enfance, les dimensions d'anxiété, d'impulsivité et la dimension négative, ont obtenu des accords inter-juges forts ; la dimension cognitive un accord inter-juge presque parfait.

Les accords inter-juges concernant la période de l'adolescence apparaissent moins élevés que pour la période de l'enfance.

4. Pourcentage de concordance brute

Une autre manière d'investiguer la fidélité inter juges est de calculer le pourcentage de concordance brute.

Cette mesure simple constitue **le rapport entre le nombre d'observations concordantes sur le nombre d'observations totales**, pour deux ou plusieurs juges, exprimé en pourcentage.

a) Pourcentage de concordance brute des items dans chaque dimension

Les tableaux ci-dessous présentent le pourcentage de concordance brute obtenu pour chacun des items, pour la période de l'enfance.

ENFANCE

▪ Dimension d'impulsivité :

passages à l'acte auto/hétéro-agressif	86%
troubles du comportement NS	88%
autres	78%

▪ Dimension d'anxiété :

anxiété/angoisse de séparation	80%
éléments phobiques	92%
éléments obsessionnels	96%

▪ **Dimension physique :**

énurésie	82%
anorexie	92%
troubles du sommeil	76%
manifestations psychosomatiques	96%

▪ **Dimension cognitive :**

QI bas	94%
retard du développement moteur	88%
retard de langage	92%
autres	90%
troubles de la latéralisation	90%

▪ **Dimension négative :**

retrait	90%
difficultés relationnelles	86%
stéréotypies	90%

▪ **Dimension positive/de discordance :**

bizarries	96%
fabulations	94%

▪ **Dimension thymique :**

items dépressifs	92%
idées suicidaires	100%
fluctuations thymiques	98%

Les tableaux ci-dessous présentent le pourcentage de concordance brute obtenu pour chacun des items, pour la période de l'adolescence.

ADOLESCENCE

▪ Dimension d'impulsivité :

passages à l'acte auto/hétéro-agressif	72%
comportement délictueux	96%
autres	66%

▪ Dimension d'anxiété :

anxiété/attaque de panique	70%
éléments obsessionnels	82%
éléments phobiques	94%

▪ Dimension thymique :

items dépressifs	76%
idées suicidaires	74%
fluctuations thymiques	78%

▪ Dimension cognitive :

trouble de l'attention / concentration	82%
difficultés d'abstraction	92%
lenteur idéative	80%
autres	80%

▪ **Dimension négative :**

retrait	74%
froideur affective	82%
avolition	70%
alogie	72%

▪ **Dimension positive/de discordance :**

discordance	84%
hallucinations	70%
persécution	74%
idées délirantes	64%

▪ **Dimension physique :**

perte de poids/anorexie	74%
troubles du sommeil	68%
incurie	84%
manifestations psychosomatiques	92%

b) Analyse

Cette mesure simple n'est que trop peu souvent utilisée car elle ne tient pas compte du hasard et ne peut répondre seule à l'exigence de la méthode scientifique. Ce n'est pas à proprement parler un test statistique.

Néanmoins elle présente l'avantage de ne pas dépendre de la prévalence et de la distribution des effectifs marginaux (tableaux de contingence creux) et s'avère facilement et rapidement réalisable. C'est pourquoi certains auteurs conseillent de l'utiliser en complément de la mesure du Kappa de Cohen [63]. Avec la mesure de la concordance brute en pourcentage, on peut comparer la fidélité du relevé même pour les items qui sont rares dans la population étudiée (c'est le cas par exemple pour les éléments obsessionnels pendant l'enfance et les éléments phobiques à l'adolescence).

Elle prend en compte une concordance globale correspondant à un accord sur la présence mais également sur l'absence de signes retrouvés.

Cette concordance brute fait apparaître des pourcentages d'accord entre les juges qui sont tous supérieurs à 70 % sur les différents items de l'IMEC utilisé.

F. Limites de la mesure de la fidélité inter-juges

Il est important de noter certains biais et difficultés rencontrés dans la mesure de la reproductibilité.

- Concernant les comportements des juges:

Nous avons observé que le juge 1 relève significativement moins de symptômes que le juge 2.

Dans cette étude, les juges ont été sélectionnés sur les critères de leur formation psychiatrique commune et sur le fait qu'ils participaient à l'étude initiale et disposaient d'une même information concernant le projet.

Le juge 1 est celui qui a décrit le contenu des items de l'IMEC ; en conséquence, nous pouvons lui supposer une aisance plus accrue à la discrimination de ces items et à l'utilisation globale de l'outil, expliquant cette différence de comportement.

- Concernant l'instrument de mesure :

Il faut noter que les deux juges n'ont pas utilisé strictement la même version. Ces différences concernaient deux points distincts :

- La grille comprenait dans sa première version, l'analyse chez l'adolescent et chez l'enfant des mêmes dimensions sauf en ce qui concerne la dimension thymique. Pour une meilleure

homogénéité, elle a donc été modifiée avec ajout de la dimension thymique pour l'enfant. Le reste de l'instrument n'a pas été modifié.

Nous avons tout de même choisi d'étudier la reproductibilité inter-juges des items de la dimension thymique, en considérant que le juge 1 n'avait pas relevé de symptômes thymiques pendant l'enfance. Cependant cette situation particulière crée une configuration de tableau de contingence entièrement creux pour le juge 1 et ne permet donc pas une estimation du coefficient Kappa, seuls les taux de concordance brute en pourcentage peuvent être examinés.

- Le manuel de passation a été formalisé pour la mesure effectuée par le juge 2, afin de préciser le contenu des items; ce qui peut interférer dans son comportement. On observe d'ailleurs une différence de comportement entre les deux juges : le juge 1 relevant significativement moins de symptômes que le juge 2.

- Concernant la méthode statistique employée :

La mesure de la subjectivité du recueil s'effectue par le calcul du coefficient Kappa de Cohen. L'interprétation de la valeur retrouvée n'est cependant pas chose aisée. Sans doute n'a-t-elle pas été conçue pour être utilisée dans le champ de l'investigation psychiatrique où comme nous le présentons dans ce travail, la place de la subjectivité y est très singulière.

Comme nous l'avons vu précédemment, selon Fleiss [46], pour déterminer l'intervalle de confiance du Kappa, avec deux modalités de jugements (présence/absence), la taille de l'échantillon devrait être $n > 64$.

Un effectif de 50 dossiers, ne permet donc pas de calculer avec suffisamment de certitude la significativité du coefficient Kappa, et il conviendrait d'augmenter le nombre de patients inclus dans l'étude.

En conclusion, l'utilisation du coefficient Kappa comme mesure de la reproductibilité est d'interprétation difficile.

Son utilisation dans une étude des phases précoces de la schizophrénie doit être très prudente car on investigue des signes peu fréquents et la prévalence diminue la valeur du Kappa attendu.

G. Discussion

1. Interprétation de la fidélité inter-juges

a) Remarques générales

Selon le barème proposé par Landis et Koch, on accepte comme mesures reproductibles celles pour lesquelles les valeurs de Kappa $>0,60$. Ce barème a été établi, lorsque le coefficient Kappa a été créé et était alors fortement utilisé dans les sciences dures comme en biologie par exemple.

Cependant comme nous l'avons vu précédemment, il n'existe pas de valeur seuil de Kappa pour laquelle, la concordance est bonne en soi. L'interprétation du Kappa est arbitraire et dépend de l'objectif de son utilisation, du type d'étude réalisée et de la nature de l'objet investigué.

Dès lors que la valeur Kappa est supérieure à 0, l'accord entre les juges, selon la définition, est supérieur à celui que produirait le hasard. Les valeurs de Kappa <0 , définissent un accord qui est inférieur à celui que produirait le hasard, mais ne signifient pas un désaccord.

Dans un cadre de recherche, certains chercheurs définissent donc un accord entre juges acceptable à partir d'une valeur de Kappa >0 , comme nous l'avons retrouvé dans cette étude neurologique sur les traumatisés crâniens [64].

Dans notre étude, il faut prendre en considération certaines particularités :

- L'impression clinique est essentielle pour permettre l'utilisation de la grille, car les symptômes relevés sont pour la plupart de nature subjective : ainsi l'humeur, l'anxiété, le contenu psychique, les idées, font appel à la inter-subjectivité dans leur mesure ;

- L'utilisation d'un recueil rétrospectif constitue une difficulté majeure car elle impose au juge de n'utiliser que les éléments du dossier sans pouvoir investiguer ailleurs, sans pouvoir faire préciser au patient ou à la famille l'exactitude du symptôme ;
- Le relevé des symptômes s'effectue selon une loi du tout ou rien, laissant peu de marge de manœuvre ;
- Le matériel utilisé est constitué de dossiers en perpétuel mouvement, certaines parties ont pu être égarées, et ainsi les deux juges n'ont pas strictement étudié les mêmes contenus de dossiers, cet effet n'est pas mesurable mais constitue un biais abaissant la concordance attendue ;
- Les dossiers pédopsychiatriques sont en grande partie des données manuscrites, constituant parfois une confusion dans leur lecture. Ce ne sont pas des dossiers de recherche mais de clinique.
- Le nombre de catégories de jugement utilisées dans l'instrument est élevé : en effet, nous relevons **25 items différents** lors de l'adolescence et **23 items différents** lors de l'enfance, soit un **total de 48 items**.

Cela a pour conséquence de **diminuer la reproductibilité** attendue par la mesure du coefficient Kappa.

Selon F. Santos, ingénieur d'études au CNRS, « l'interprétation du Kappa de Cohen doit être modulée par le nombre de stades possibles » [54]. En effet, le Kappa aura toujours tendance à être plus faible pour un grand nombre de stades possibles, que pour seulement deux stades à attribuer : il est plus facile de se mettre d'accord sur « Plutôt Noir / Plutôt Blanc » que sur toute une palette comportant 5000 gris !

Ainsi, un **Kappa = 0,40** pourra être considéré comme très médiocre si les 2 juges avaient seulement à choisir entre deux scores A et B, **mais pourra être perçu comme relativement honorable** [54] **s'ils doivent choisir parmi plus de 20 catégories différentes** (les items) pour chaque période considérée.

- En revanche les catégories utilisées ne sont pas exclusives c'est à dire qu'un juge peut choisir plusieurs catégories de jugement, rendant plus probable la concordance aléatoire (superposition de certaines réponses par le seul fait du hasard) et diminuant la concordance maximale attendue.

- La table de Landis et Koch ne fournit qu'un guide de lecture, à compléter par l'expertise du praticien. Comme nous le signalent nombre de statisticiens dont Byrt et Bishop [66] ou bien Lantz et Nebenzahl [67] ; le paramètre Kappa présente des limites dans son utilisation et doit donc être manié avec prudence. Nous allons étudier ses limites en décrivant certains biais et la manière dont ils influent sur la valeur de Kappa. Notamment, dans certains cas, on peut retrouver des résultats paradoxaux c'est à dire qu'il peut exister une concordance forte associée à un Kappa faible. Ce paradoxe est la conséquence de deux caractéristiques du paramètre statistique Kappa et a été décrit par Feinstein et Cicchetti [68]. Ce phénomène, repose sur le fait que :
 - la valeur de Kappa dépend de la prévalence du signe dans la population étudiée, nous devons considérer que plus le signe est fréquent dans la population, plus le Kappa que nous devons attendre est grand. Cela signifie également de manière réciproque, que les faibles valeurs de Kappa ne signifient pas forcément un accord faible entre les juges, pour les critères dont l'occurrence est rare [69, 70, 71].

 - la valeur de Kappa est influencée par la distribution du signe dans le tableau de contingence. Pour des proportions d'accord identiques, Kappa sera plus grand pour des effectifs marginaux symétriques et/ou déséquilibrés. La compréhension de ce phénomène dépasse le cadre de notre étude et fait appel à des connaissances statistiques plus approfondies qui n'ont pas été utilisées dans notre travail.

L'accord inter-juges est un phénomène complexe, nous avons décidé de l'aborder sur plusieurs niveaux : par l'estimation du coefficient Kappa global, puis selon le regroupement dimensionnel proposé et enfin selon les coefficients Kappa catégoriels analysant chacun des items de l'IMEC.

b) Seuil du coefficient Kappa retenu dans notre analyse

Etant donné la particularité de l'objet psychiatrique et le nombre élevé de catégories utilisées; nous avons choisi, de définir **un seuil Kappa > 0,40** pour accepter une concordance recevable dans le relevé des items de l'IMEC.

Une adaptation du barème de référence est suggérée dans des articles récents pour les études du domaine psychiatrique [53]. Ce seuil, qui témoigne selon Landis et Koch d'un accord au minimum considéré comme modéré, est utilisé dans des études scientifiques récentes [65] discutant la fidélité inter-juges. Il nous a paru approprié dans ce cadre de recherche afin de distinguer les items les moins reproductibles et envisager par la suite une modification de l'instrument.

Il constitue dans notre étude, le « minimum acceptable » pour conserver un item de l'IMEC :

- pour les valeurs de Kappa <0,40, la suppression de l'item est envisagée ;
- pour les valeurs de Kappa entre 0,40 et 0,60 ; une modification du contenu peut être proposée, pour reformuler la description de l'item ;
- pour les valeurs de Kappa entre 0,60 et 0,80 ; des précisions dans la description peuvent être apportées, afin de diminuer la part d'interprétation laissée au juge ;
- pour les valeurs de Kappa >0,80 ; l'item est conservé tel qu'il est formulé dans l'IMEC.

c) Coefficient Kappa global de l'IMEC

La moyenne pondérée des coefficients Kappa catégoriels permet d'obtenir l'estimation d'un coefficient Kappa global ayant pour valeur Kappa = 0,43.

Selon la table de Landis et Koch, on peut donc définir qu'il existe un **accord global inter-juges modéré** pour l'utilisation de l'IMEC.

d) Coefficients Kappa par dimension

Pendant l'adolescence, toutes les dimensions de l'IMEC apparaissent reproductibles, avec un même accord inter-juges considéré comme modéré ; en dehors de la dimension thymique (accord faible; Kappa = 0,31).

Pendant l'enfance, toutes les dimensions apparaissent reproductibles, avec un **accord inter-juges** allant de **modéré** pour les dimensions physique et positive ; un **accord fort** pour les dimensions impulsivité, anxiété et négative ; et un **accord presque parfait** pour la dimension cognitive.

Ces résultats montrent une reproductibilité des données recueillies concernant l'utilisation des dimensions. Certaines remarques méritent d'être faites :

- Nous pouvions nous attendre à des valeurs de Kappa plus élevées pour deux raisons :
 - d'une part, l'étude initiale suggérait une continuité des dimensions présentées entre l'enfance et l'adolescence, ce qui faisait envisager une stabilité meilleure des résultats dans cette conception dimensionnelle ;
 - d'autre part, chaque dimension a été considérée comme validée lorsqu'au moins un des items la constituant était retrouvé par le juge, ce qui constitue un critère de reproductibilité moins exigeant.

- Nous observons un accord très élevé pour la dimension cognitive pendant l'enfance. Les signes investigués par cette dimension sont retrouvés préférentiellement dans les dossiers des enfants ayant présenté un trouble envahissant du développement. Ce sont des dossiers volumineux où nous disposons sans doute d'un nombre d'informations plus important, ce qui permet d'augmenter la fidélité du recueil.

e) Coefficients Kappa catégoriels

Selon le seuil de 0,40 déterminé pour notre étude de la reproductibilité des résultats et la table de Landis et Koch :

(1) Les items de l'instrument de mesure utilisé que l'on retient comme **reproductibles**, classés par valeur du coefficient Kappa, sont :

Pendant la période de l'enfance :

- l'item « retard de langage » qui bénéficie d'un accord considéré comme presque parfait ou très bon (Kappa = 0,80-1,00).
- les items disposant d'un accord considéré comme fort (Kappa = 0,60-0,80) :
 - « manifestations psychosomatiques »
 - « QI bas »
 - « retard du développement moteur »
 - « autres troubles cognitifs »
 - « retrait » ;
- les items avec un accord considéré comme modéré (Kappa = 0,40-0,60) :
 - « passages à l'acte auto ou hétéro agressif »
 - « troubles du comportement non spécifiés »
 - « autres troubles du comportement »
 - « anxiété /angoisse de séparation »
 - « éléments phobiques »
 - « énurésie »
 - « anorexie »
 - « troubles de la latéralisation »
 - « stéréotypies »
 - « bizarreries »
 - « fabulations » ;

Pendant la période de l'adolescence :

- l'item « comportements délictueux » dispose d'un accord considéré comme presque parfait (Kappa = 0,80-1,00).

- l'item « discordance » dispose d'un accord considéré comme fort (Kappa = 0,60-0,80) ;
- les items avec un accord considéré comme modéré (Kappa = 0,40-0,60) :
 - « passages à l'acte auto ou hétéro agressifs »
 - « anxiété ou attaque de panique »
 - « éléments obsessionnels »
 - « items dépressifs »
 - « trouble de l'attention/concentration »
 - « difficultés d'abstraction »
 - « lenteur idéative »
 - « retrait »
 - « froideur affective »
 - « avolition »
 - « persécution »
 - « incurie »
 - « manifestations psychosomatiques ».

(2) Les items que l'on retient comme **non reproductibles** sont, selon ce même barème, ceux pour lesquels nous avons relevé un coefficient Kappa <0,40, signifiant un accord difficilement acceptable car au mieux qualifié de faible :

Pendant la période de l'enfance, aucun item n'est apparu non reproductible.

Pendant la période de l'adolescence, les items suivant :

- « autres troubles du comportement »
- « idées suicidaires »
- « troubles du sommeil »
- « perte de poids/anorexie ».

(3) Le calcul du coefficient Kappa n'est pas significatif pour les autres items, sans doute en raison d'un trop faible effectif de l'échantillon :

Pendant la période de l'enfance, les items suivant :

- « obsessions »
- « troubles du sommeil »
- « difficultés relationnelles »

Pendant la période de l'adolescence, les items suivant :

- « phobies »
- « fluctuations thymiques »
- « autres troubles cognitifs »
- « alogie »
- « hallucinations »
- « délire »

Pour les items cités ci-dessus, nous ne pouvons nous prononcer sur la fidélité inter-juges à partir du coefficient Kappa de Cohen. Nous utiliserons donc en complément la mesure du pourcentage de concordance brute pour conclure sur la reproductibilité des résultats.

2. Synthèse

Si la mesure globale de la fidélité inter-juge de l'IMEC montre une reproductibilité acceptable des résultats, les accords inter-juges sur de nombreux items nous sont apparus décevants.

En effet, peut être naïvement, avant la réalisation de cette étude, nous attendions une valeur entre 0,60 et 1 pour le coefficient Kappa, qui aurait signifié un accord fort à presque parfait entre les deux juges.

Comment expliquer les taux de fidélité inter-juges observés ?

Les résultats pour lesquels le coefficient Kappa est le moins bon sont en fait des résultats non significatifs sur le plan statistique.

Pour la période de l'enfance, cette situation peut s'expliquer par la taille trop faible de l'échantillon. Sur les 50 dossiers, seulement 25 ont un suivi pendant l'enfance réduisant l'échantillon de moitié.

Les signes qui sont objectivés dans chaque item ne le sont que sur trop peu de patients rendant la mesure du kappa non significative.

Pour la période de l'adolescence, des difficultés dans la mesure même des données peut s'expliquer par :

- la complexité de l'utilisation du coefficient Kappa de Cohen et la faible prévalence des signes recherchés : là aussi, les signes qui sont objectivés ne le sont que sur peu de patients rendant la mesure du kappa non significative ;
- la difficulté de la distinction entre prodromes et début de la maladie : les deux juges ont probablement eu des interprétations différentes dans la définition d'un épisode psychotique caractérisé témoignant du début de la schizophrénie. Sans doute existe-t-il un continuum entre les signes précoces et les signes d'une schizophrénie caractérisée. Par conséquent, la discrimination de signes présentés « avant » et « après » l'apparition de la maladie qui a été effectuée dans cette étude a pu être source d'une grande part de la variabilité des mesures donc des désaccords entre les juges.

3. Modifications de l'IMEC

Cette étude de la fidélité inter-juges a eu plusieurs impacts.

Elle nous a permis dans un premier temps de classer les items de l'IMEC selon leur subjectivité par rapport aux résultats obtenus par les différents tests.

D'autre part, elle a fait envisager des sources de désaccord et nous incite à réformer en conséquence l'instrument utilisé dans ce cadre de recherche.

Suite à ce premier travail d'évaluation de la fidélité inter-juges montrant de nombreuses non congruences sur de nombreux items, l'équipe de recherche du Professeur Kabuth s'est réunie de nombreuses fois et nous proposons un certain nombre de modifications :

a) *Au niveau des modalités de réponse*

L'existence de 2 modalités « présent/absent » apparaît trop restreinte.

En effet, certains dossiers ne permettent pas au juge de déterminer si la présence du signe existe mais pour autant ne lui permettent pas non plus de déterminer que ce signe est absent. Ce phénomène a pu être une source de confusion dans la mesure par les juges utilisant l'IMEC.

Nous proposons donc d'introduire une troisième modalité de jugement des items appelée « Non déterminable ».

Il existe également une part d'interprétation dans l'utilisation de cet instrument qu'il est nécessaire de pouvoir mesurer.

Nous proposons d'élargir le choix des modalités de jugement de chaque item, en introduisant la notion de déduction des informations retenues.

Les modalités de jugement pourraient être ainsi distinguées selon la démarche effectuée par l'utilisateur de l'IMEC pour effectuer la cotation. Nous retiendrons au total 5 modalités de jugement, décrites comme ci-dessous pour chaque item:

Présence :

- cocher « **e** » (pour « écrite ») lorsque le symptôme est écrit/décrit dans le dossier ;
- cocher « **d** » (pour « déduite ») lorsque le symptôme n'est pas clairement écrit dans les dossiers mais que les informations contenues sont jugées suffisantes pour déduire sa présence.

Absence :

- cocher « **e** » lorsque l'absence du symptôme est stipulée dans le dossier (par exemple : *pas d'angoisse*) ;
- cocher « **d** » lorsque l'absence du symptôme n'est pas clairement stipulée dans le dossier mais que les informations y figurant sont suffisantes pour la déduire.

ND : non déterminable, c'est-à-dire que les éléments du dossier ne permettent pas de déterminer l'absence ou la présence du symptôme.

b) Au niveau de la structure

Nous proposons de diminuer le nombre de catégories comme le proposent Maclure et Willet [70]; ce qui permettra d'augmenter la significativité du test Kappa, surtout pour les items de faible prévalence dans la population étudiée, et procéder à un regroupement des items lorsque cela est possible.

c) Au niveau du contenu des items

Nous avons étudié les périodes de l'enfance et de l'adolescence séparément.

Nous discuterons d'abord le contenu des items pour lesquels le coefficient Kappa est non significatif.

Il est alors impossible de conclure sur la fidélité inter-juges à partir du seul coefficient Kappa et nous avons donc pris en considération le **complément d'information** apporté par le pourcentage de concordance brute.

Dans cette configuration, qui ne tient pas compte du hasard, il n'existe pas de table de référence comme c'est le cas pour le coefficient Kappa. Nous avons choisi de fixer l'objectif d'un taux de concordance brute $>80\%$, afin de retenir une mesure comme reproductible lorsque le coefficient Kappa est non significatif.

Cette situation où l'on retrouve un coefficient Kappa faible, mais de manière non statistiquement significative, et un pourcentage de concordance brute élevé s'explique par une prévalence faible du signe recherché dans la population étudiée. L'accord se situe alors plutôt dans l'absence du signe recherché mais cet aspect n'est pas pris en compte par le coefficient Kappa.

Par la suite, des modifications des items où le Kappa est entre 0,4 et 0,6 sont proposées.

Enfin, la suppression des items où le Kappa est inférieur à 0,4 est envisagée.

Dans ces deux derniers cas, le taux de concordance brute n'apporte pas d'élément contraire à la conclusion sur la fidélité inter-juges obtenue par l'estimation du coefficient Kappa, car tous les taux de concordance brute observés se situent en dessous de 80% et le coefficient Kappa apparaît significatif.

(1) Période de l'enfance

(a) *Cas des items où le kappa est non significatif*

▪ **Dimension d'anxiété**

Eléments obsessionnels : le coefficient Kappa est négatif (-0,02), et indique un désaccord entre les juges cependant le pourcentage de concordance brute est très élevé (96%).

Nous proposons de conserver l'item, la faible valeur de l'estimation du Kappa s'expliquant par une prévalence très faible du signe (effectif mesuré par les deux juges est égal à 1).

Une nouvelle mesure pourrait être envisagée après augmentation de la taille de l'échantillon.

▪ **Dimension physique**

Troubles du sommeil : le coefficient Kappa est faible (0.12) et le pourcentage de concordance brute est inférieur au seuil retenu pour la reproductibilité des résultats mais s'en rapproche (76%). Les effectifs observés sont modérés (10 par le juge 1 et 6 par le juge 2).

La suppression de l'item paraît difficile, car il s'agit d'un symptôme fréquemment retrouvé en pédopsychiatrie. Sa description dans les dossiers est apparue trop indéfinie et mériterait des précisions.

▪ **Dimension négative**

Difficultés relationnelles : le coefficient Kappa est faible (0.15) et le pourcentage de concordance brute est élevé (86%). Le signe est peu fréquent dans l'échantillon de population observée (le juge 1 retrouve un effectif n=3, le juge 2 un effectif n=6).

L'analyse du contenu de l'item montre des imprécisions dans la définition des difficultés relationnelles (difficultés à entrer en relation avec le groupe de pairs, troubles de l'adaptation à l'école, terme « difficultés relationnelles » peut être noté dans le dossier sans plus de précisions).

Peut être existe-t-il un abus de langage pour la période de l'enfance ; et une tendance à inclure dans les difficultés relationnelles un large spectre de symptômes comme la timidité, l'inhibition, l'immaturité... ?

Nous décidons de proposer un remaniement des items de la dimension négative : le retrait (Kappa =0,72) est conservé, tout comme les stéréotypies motrices (Kappa =0,50).

Les troubles de l'adaptation à l'école et les difficultés à entrer en relation avec le groupe de pairs ainsi que les difficultés relationnelles mentionnées sans plus de précision sont supprimés de la

description de l'item et ne sont pas inclus par ailleurs dans la dimension négative car ils paraissent être trop sujets à l'interprétation du cotateur.

(b) Cas des items où le kappa est compris entre 0,40 et 0,60

▪ **Dimension positive/ de discordance**

Bizareries : le coefficient Kappa a une valeur de 0,48 ; l'effectif relevé est faible (n=2) pour les 2 juges diminuant la valeur attendue de Kappa.

Ces symptômes sont souvent retrouvés chez les enfants présentant un trouble envahissant du développement et peuvent avoir une valeur annonciatrice importante dans l'occurrence des troubles du spectre schizophrénique ultérieure [55].

L'item est donc conservé et nous proposons des précisions dans sa définition ci-dessous :

Bizareries du comportement/ des propos: terme précisé et décrit dans les dossiers en tant que tel, éléments pouvant faire évoquer une discordance à la lecture des dossiers : incohérence des propos, maniérisme gestuel (marche sur pointe des pieds...) et verbal (néologisme).

Fabulations : la même démarche est effectuée pour cet item où le coefficient Kappa est de 0,55 et la prévalence faible (effectif n=2 pour le juge 1, n=6 pour le juge 2).

Ce signe n'a pas été suffisamment décrit dans l'IMEC et laisse place à l'interprétation du cotateur. L'item est conservé en proposant la définition suivante :

Fabulations envahissantes, imaginaire envahissant: le terme fabulation doit être précisé dans les dossiers en tant que tel, cela peut prendre de telle proportion jusqu'à suspecter des idéations délirantes. Le caractère «envahissant» est précisé afin de ne pas prendre en compte les fabulations « normales » de l'enfant, ceci pouvant correspondre à une source de confusion entre les juges.

▪ **Dimension d'anxiété**

Eléments phobiques : le coefficient Kappa vaut 0,56 ; le signe est peu fréquent dans l'échantillon observé et peut expliquer une valeur modérée du Kappa.

L'item peut être conservé en apportant des précisions dans la définition.

On peut notamment indiquer de ne pas y inclure les peurs qui peuvent être normales dans l'enfance et insister sur le caractère phobique. Il faut enlever le terme « peur » de la définition car celui-ci apporte une confusion.

Anxiété/angoisse de séparation : le coefficient Kappa a une valeur modérée de 0,46 ; le symptôme est relativement fréquent (n=9 et 15). Nous proposons de conserver l'item en modifiant sa définition concernant la notion d'angoisse de séparation. Nous retenons uniquement les difficultés à se séparer d'une figure d'attachement quand cela est précisé dans les dossiers, les autres formes possibles d'expression ne sont pas retenues car l'interprétation peut être trop importante.

- **Dimension cognitive**

Troubles de la latéralisation : le coefficient Kappa vaut 0,56. Nous proposons de conserver ce signe, en l'incluant dans un item plus général « troubles psychomoteurs ». Il est possible d'avoir compris « par excès » dans les troubles de latéralisation d'autres anomalies motrices, faute d'item spécifique. Nous proposons des modifications de l'item « retard du développement moteur » redéfini en deux items permettant de faire la distinction entre: d'une part, le retard (des acquisitions, de la marche, du graphisme...) et d'autre part, les troubles psychomoteurs sans notion de retard (difficultés d'orientation spatio-temporelle, maladresse gestuelle, dyspraxie, trouble de la motricité fine et troubles de latéralisation).

- **Dimension négative**

Stéréotypies : la valeur du coefficient Kappa est de 0,50 ; le symptôme est peu fréquent dans l'échantillon étudié (n=3 et 8). Il s'agit d'un symptôme important car il concerne des enfants présentant un trouble envahissant du développement et le lien avec une schizophrénie ultérieure est largement évoqué dans la littérature scientifique [55].

Nous proposons de simples précisions visant à diminuer l'interprétation possible.

Stéréotypies : stéréotypies motrices ou verbales, comportement stéréotypé et écholalie sont incluses lorsque les termes sont mentionnés dans les dossiers.

▪ **Dimension comportementale**

Passages à l'acte auto/hétéro-agressifs : le coefficient Kappa a été mesuré à 0,52. Le signe est relativement fréquent (n=5 et n=12) dans l'échantillon.

Le symptôme est conservé mais nous proposons de le scinder en deux. Les passages à l'acte auto-agressifs sont conservés dans un item individualisé et l'intentionnalité suicidaire quand elle est retrouvée doit être précisée.

Les passages à l'acte hétéro-agressifs peuvent être classés dans un nouvel item intitulé « manifestations externalisées », plus parlant au niveau de la séméiologie pédopsychiatrique et qui devrait limiter la confusion possible.

Troubles du comportement non spécifiés : le coefficient Kappa a une valeur estimée à 0,43. Nous proposons d'inclure ces signes dans l'item créé intitulé « manifestations externalisées » décrites ci-dessus.

Autres : le coefficient Kappa mesuré vaut 0,46. Nous proposons de conserver les différents signes constituant cet item et de les inclure également dans les « manifestations externalisées ».

Au final, dans cette dimension comportementale, 2 items sont donc différenciés :

- passage acte auto-agressif (avec ou sans intentionnalité suicidaire)
- manifestations externalisées (troubles du comportement non précisés, impulsivité, intolérance à la frustration, opposition, irritabilité, agitation psychomotrice, fugues, crises clastiques, colères fréquentes).

▪ **Dimension physique**

Énurésie : le coefficient Kappa est de 0,46 ; le symptôme est relativement fréquent (n=10 et n=11). Pour ce symptôme, le lien avec la schizophrénie est bien défini dans la littérature où il est décrit comme le témoin d'anomalies développementales [55].

L'item est conservé en apportant des précisions : l'énurésie peut être distinguée en énurésie primaire ou secondaire.

Anorexie/perte de poids : le coefficient Kappa a une valeur de 0,56 ; le symptôme est peu fréquent (n=4 et n=6). Nous proposons une précision supplémentaire de l'item. En effet, les juges ont tendance à y inclure d'autres troubles du comportement alimentaires, faute d'items

spécifiques. Nous proposons donc une définition plus explicite incluant d'autres troubles du comportement alimentaire sous la dénomination suivante :

Trouble du comportement alimentaire : anorexie/perte de poids non organique, hyperphagie, potomanie, boulimie.

(c) Cas des items où le coefficient Kappa est inférieur à 0,4

Aucun item durant la période de l'enfance ne correspond à cette situation.

(2) Période de l'adolescence

(a) Cas des items où le coefficient Kappa est non significatif

Nous envisageons, comme pour la période de l'enfance, de considérer le complément d'information qu'apporte le pourcentage de concordance brute, pour proposer des pistes de réflexion dans la modification du contenu des items.

▪ Dimension positive

Hallucinations : le coefficient Kappa est estimé à 0,05 et le pourcentage de concordance brute à 70%.

Idées délirantes : le coefficient Kappa est estimé à 0,17 et le pourcentage de concordance brute à 64%.

Ces 2 signes ont une valeur sémiologique annonciatrice de la schizophrénie très importante [55]. Nous ne proposons pas de suppression pour ces items même si la valeur de Kappa est faible; il semble que les discordances observées soient en rapport avec une définition imprécise de la période pré-schizophrénique dans les dossiers et de la difficulté à distinguer le « début » du premier épisode psychotique caractérisé de la « fin » de la période prodromique.

Ainsi, les hallucinations et les idées délirantes ont pu être comptabilisées par un juge comme étant des prodromes et par l'autre juge comme appartenant aux symptômes du premier épisode psychotique. Ceci est d'autant plus vrai pour les débuts insidieux, renforçant la confusion possible.

Des précisions doivent être apportées à l'instrument pour délimiter avec plus de rigueur les prodromes des symptômes du premier épisode psychotique.

- **Dimension cognitive**

Alogie : le coefficient Kappa est de 0,18 et le pourcentage de concordance brute est de 72%. Les effectifs relevés sont très différents selon le juge (n=5 pour le juge 1 et n=15 pour le juge 2).

Après discussion entre les juges, il s'avère que l'item « alogie » a été mal compris par le juge 2, qui l'envisageait comme un trouble du système logique de raisonnement au lieu d'un trouble du langage avec pauvreté du discours ou barrages (comme décrit dans le manuel de passation). Cet événement montre à quel point, les séances de familiarisation avec l'outil sont nécessaires avant son utilisation pour un recueil standardisé des données.

Nous retenons cet incident comme source de désaccord et proposons de supprimer l'item « alogie » individualisé dans l'instrument initial pour inclure certains symptômes le constituant (les barrages) dans l'item « discordance » de la dimension positive avec la définition ci-dessous :

Discordance : bizarrerie du comportement, du contact, des propos, maniérisme, rires immotivés, diffluence de la pensée, déréalisation/dépersonnalisation, dissociation, hermétisme, ambivalence et barrages.

- **Dimension d'anxiété**

Phobie : le coefficient Kappa (-0,03) montre un désaccord des juges. Le pourcentage de concordance brute (72%) est inférieur au seuil de 80%.

Aucune source de confusion n'a pu être évoquée. Nous pouvons donc envisager une suppression de cet item.

- **Dimension thymique**

Fluctuations thymiques : le coefficient Kappa est très faible (0,03) et le pourcentage de concordance brute (78%) se rapproche du seuil fixé à 80%.

Nous proposons des précisions dans la définition de l'item.

Fluctuations de l'humeur : symptômes qui se réfèrent à la labilité thymique, versatilité de l'humeur, élation de l'humeur sur un temps restreint. Nous proposons une distinction avec

« l'excitation psychomotrice » (avec note thymique): logorrhée, coq à l'âne, associations par assonance évoquant des éléments du registre maniaque ou hypomaniaque, et l'individualisons. Le fait d'avoir initialement regroupé les fluctuations thymiques et l'excitation psychomotrice peut avoir été source de confusion.

▪ **Dimension cognitive**

Autres : le coefficient Kappa est faible (0,05) et le pourcentage de concordance brute (80%) se situe à la limite exacte du seuil de reproductibilité fixé.

Cette catégorie apparaît trop imprécise et regroupe des symptômes trop variés (troubles de l'attention, de la concentration, difficultés d'abstraction, lenteur idéative, autres :maladresse gestuelle et troubles des fonctions instrumentales).

Il est probable que l'utilisateur de l'IMEC ait tendance à vouloir y inclure d'autres symptômes ne correspondant à aucun item. Nous proposons sa suppression afin d'améliorer la fidélité inter-juges du relevé.

(b) Cas des items où le coefficient Kappa est compris entre 0,4 et 0,6

▪ **Dimension comportementale d'impulsivité**

Passages à l'acte auto/hétéro-agressifs : le coefficient Kappa est mesuré à 0,44. Le symptôme est très fréquemment observé dans l'échantillon (n=21 et n=25).

Nous proposons donc de le conserver et apportons des précisions sur l'intentionnalité suicidaire :

- Passages à l'acte auto-agressifs avec intentionnalité suicidaire claire (IMV, phlébotomie...)
- Passages à l'acte auto-agressifs sans intentionnalité suicidaire (scarifications...).

Nous proposons d'individualise les passages à l'acte hétéro-agressifs dans les troubles du comportement externalisés.

▪ **Dimension cognitive**

Difficulté d'abstraction : le coefficient Kappa est de 0,46. Le symptôme est peu fréquent (n= 3 et n=5). Au lieu d'envisager des précisions ou une modifications du contenu de l'item, nous proposons sa suppression car il s'agit d'un terme peu employé en routine et pouvant prêter à confusion.

Troubles de l'attention/concentration : le coefficient Kappa a une valeur modérée de 0,42. Nous proposons de conserver cet item qui est fréquemment retrouvé (n=10 et n=9).

Lenteur idéative : le coefficient Kappa est également de 0,42. Nous envisageons ici la suppression de l'item qui est trop imprécis et peut être le reflet d'autres signes déjà cotés ailleurs dans l'IMEC : QI bas, barrage, trouble du cours de la pensée, inhibition...

▪ **Dimension d'anxiété**

Anxiété/attaques de panique : le coefficient Kappa vaut 0,40. Le symptôme est très fréquent dans la population étudiée (n=25 et n=32).

L'item peut être conservé en apportant des précisions supplémentaires notamment en ajoutant dans la définition : les angoisses massives et les angoisses de morcellement qui ont pu être incluses auparavant dans l'item « discordance » seul.

Éléments obsessionnels : le coefficient Kappa est de 0,53. Nous ne proposons pas de modification de cet item.

▪ **Dimension négative**

Retrait : le coefficient Kappa a une valeur de 0,45. L'item est conservé en apportant des précisions. Le retrait se réfère également aux termes « isolement », « inhibition majeure », « renfermé » dans le sens d'un repli social. Cette précision apparaît importante afin de ne pas y inclure des symptômes « par excès ».

Froider affective : le coefficient Kappa vaut 0,53. L'item peut être conservé.

Avolition : le coefficient Kappa est de 0,42. L'item peut être précisé: correspond au désintérêt, désinvestissement (y compris à l'école), diminution du rendement dans les différentes activités, apragmatisme, apathie. C'est-à-dire les éléments qui se réfèrent à une altération du fonctionnement habituel. Les symptômes de désinvestissement scolaire (comme l'absentéisme), témoignant d'un

désintérêt sont inclus. En revanche, les symptômes de diminution des résultats scolaires relèvent plutôt des symptômes cognitifs.

- **Dimension physique**

Incurie : le coefficient Kappa est de 0,46. Nous proposons ici de conserver l'item mais de l'inclure plutôt dans la dimension négative où il y prendra plus de « sens ».

Manifestations psychosomatiques : le coefficient Kappa est de 0,46. Nous proposons de conserver cet item.

- **Dimension thymique**

Items dépressifs : le coefficient Kappa est mesuré à 0,52. Le symptôme est très fréquemment observé dans l'échantillon (n=25 pour les 2 juges).

Nous proposons d'approfondir la définition des « items dépressifs » comme suit : ils correspondent aux termes tristesse de l'humeur, mélancolie, affects dépressifs, syndrome dépressif, dépression, ralentissement psychomoteur, idées noires, pessimisme, auto-dévalorisation, culpabilité non délirante, anhédonie retrouvés dans les dossiers.

(c) Cas des items où le coefficient Kappa est inférieur à 0,4

- **Dimension thymique**

Idées suicidaires : le coefficient Kappa a été mesuré à 0,18. Les effectifs relevés sont très différents selon le juge (n=15 pour le juge 1 et n=2 pour le juge 2).

L'analyse de la décision de cotation, en présence des juges, montre que l'intentionnalité suicidaire supposée d'un acte auto-agressif peut être comprise par un juge comme l'équivalent d'une idéation suicidaire non exprimée.

Nous proposons donc de spécifier la définition des idées suicidaires comme un contenu psychique conscient et exprimé. Les comportements auto-agressifs dans la dimension comportementale/d'impulsivité peuvent également être distingués en fonction de l'intentionnalité suicidaire qui y est associée.

▪ **Dimension physique**

Anorexie/perte de poids : le coefficient Kappa vaut 0,37 et le symptôme est fréquent (n=9 et n=18). Les mêmes remarques que pour la période de l'enfance peuvent être formulées.

Certains symptômes alimentaires non testés par l'item ont pu être inclus par excès. Nous envisageons donc des modifications majeures du contenu de l'item afin de standardiser le relevé. Les modifications sont identiques à celles proposées pour la période de l'enfance.

Troubles du sommeil : le coefficient Kappa est mesuré à 0,35 et le symptôme est très fréquent (n=16 et n=24).

Un travail de reformulation plus précise est en cours afin de distinguer les perturbations mineures et occasionnelles aux troubles caractérisés du sommeil.

▪ **Dimension comportementale /d'impulsivité**

Autres : le coefficient Kappa est mesuré à 0,31. Il s'agit d'un signe très fréquent dans l'échantillon étudié (n=26 et n=35).

La dénomination « autres » et la définition de cet item en font une catégorie « fourre-tout » sujette à inclure de nombreux symptômes non décrits ailleurs.

Nous proposons, comme pour la période de l'enfance, la suppression de cet item. Les troubles du comportement peuvent être distingués par deux catégories :

-Passages à l'acte auto-agressifs (distingués selon leur intentionnalité suicidaire)

-Troubles externalisés : comportement délictueux (délinquance, vols), passages à l'acte hétéro-agressifs, intolérance à la frustration, crises clastiques, fugues, troubles du comportement non spécifiés dans les dossiers (la mention « trouble du comportement » apparaît dans le dossier sans être détaillée), agressivité sans passage à l'acte, opposition (« comportement opposant »), agitation psychomotrice sans note thymique.

Au total, nous envisageons donc les modifications suivantes de l'IMEC:

- **pour la période de l'enfance, des précisions importantes sur le contenu définissant 8 des items et un remaniement des dimensions négative et comportementale ;**
- **pour la période de l'adolescence, des précisions importantes sur le contenu définissant 9 des items, le remaniement de la dimension comportementale comme pour la période de l'enfance et un remodelage majeur de la dimension cognitive avec la suppression de 3 items la constituant.**

H. Perspectives

Nous proposons également des compléments d'étude afin de poursuivre la recherche des signes cliniques présents lors des phases précoces de la schizophrénie.

a) Compléments d'analyse de la fidélité inter-juges

Après modification de l'instrument, il conviendrait de réaliser une nouvelle analyse de la concordance inter-juges, afin de **mettre en évidence l'amélioration de fidélité obtenue**.

Les juges devraient alors être choisis de manière **indépendante** à l'étude, en vue de favoriser leur impartialité, et la cotation des dossiers devrait être réalisée **en aveugle** de la pathologie étudiée pour diminuer le biais de mesure. Le **nombre de juges** pourrait être **augmenté** à 3, dans l'objectif d'assurer une puissance statistique supérieure de la reproductibilité des résultats.

Si possible, le nombre n de patients constituant l'échantillon devrait être augmenté afin d'obtenir $n > 16r^2$, avec r modalités de réponse.

b) Poursuite de l'étude initiale dans un but analytique

Le regroupement dimensionnel devrait être étudié par une analyse statistique, de type corrélationnelle, des différents items inclus dans une même dimension.

Nous retenons notre attention sur l'impossibilité à partir des données recueillies par l'IMEC de déterminer si les signes cliniques observés sont des manifestations prémorbides ou des prodromes spécifiques de la schizophrénie car le lien de causalité ne peut être établi dans cette étude sans groupe témoin.

L'étude initiale est une étude cas qui ne permet donc pas déterminer un lien scientifique de cause à effet. Cependant si une étude de témoins était réalisée selon le même protocole, alors l'ensemble pourrait constituer une étude d'observations à caractère analytique de type cas-témoins permettant d'établir un lien de causalité entre les signes précoces observés et l'existence de la maladie schizophrénique à l'âge adulte.

Un travail est donc envisagé pour répondre à cette question. Il est proposé de faire lire l'ensemble des dossiers sélectionnés par des juges en aveugle du statut (cas ou témoins), le choix de 2 groupes de comparaison est proposé :

- le premier, parmi la population des enfants suivis au même moment que les cas quand ils étaient enfants (sur le principe de la cohorte historique) et pour lesquels aucun diagnostic de schizophrénie n'a été porté ;
- le deuxième, parmi la population des adultes actuellement suivi au C.P.N pour un autre motif que la schizophrénie (trouble de l'humeur par exemple).

CONCLUSION

Notre travail a permis d'évaluer la fidélité inter-juges d'un instrument de mesure complexe utilisé dans une étude de recherche en pédopsychiatrie menée récemment à Nancy.

Des arguments en faveur de la reproductibilité des résultats obtenus dans cette étude initiale ont pu être avancés. Cependant nous retenons que les résultats sont éloignés de ceux auxquels nous pouvions nous attendre et nous ont amené à proposer des modifications de l'instrument utilisé.

Ces éléments nous ont fait envisager des causes possibles à cette variabilité des résultats entre les juges investiguant les mêmes dossiers. Notamment l'utilisation largement répandue du coefficient Kappa de Cohen comme mesure de la fidélité inter-juges s'est avérée délicate dans le contexte de cette étude investiguant parfois des signes peu fréquents dans la population de l'échantillon.

Par ailleurs, ce travail nous a amené à confronter la méthodologie scientifique de la recherche aux particularités de l'objet investigué en psychiatrie. Mais plutôt que de mettre en concurrence ou en opposition ces différentes approches qui ont chacune leur spécificité et leur nécessité, il nous paraît essentiel de promouvoir une autre voie qui serait celle de leur intégration dans un même cadre conceptuel.

Cette situation implique nécessairement une évolution culturelle qui est déjà largement engagée, mais qui se heurte encore dans le domaine de la recherche à des débats qui ont plus souvent conduit à la démission qu'à un véritable affrontement des problèmes. Ils concernent le statut des données objectives et plus généralement la méthodologie, ainsi que l'articulation entre clinique et recherche. Or ces deux points conditionnent de fait la possibilité de la participation des psychiatres cliniciens à la recherche, et donc en fin de compte la recherche elle-même. C'est pourquoi nous recommandons que les psychiatres puissent se former en épidémiologie et en bio-statistiques, afin d'assurer le lien entre la recherche de sens et la rigueur de la méthodologie employée.

La position de la fédération française de psychiatrie est d'ailleurs *d'une part de privilégier leur abord dans la pratique, plutôt que de façon théorique, et surtout d'éviter une fracture entre chercheurs professionnels et cliniciens.* Cette fracture aurait des conséquences très néfastes pour la recherche en psychiatrie car les moyens humains doivent correspondre à l'étendue du champ

concerné et atteindre une masse critique qui assure la possibilité même des recherches.

Références

1. Indrayan A. Elements of medical research, J. Med. Res. 2004 ; 119 (3) : 93-100.
2. Highleyman L. A guide to clinical trials. Part II: interpreting medical research. BETA 2006 ; 18 (2) : 41-47.
3. Bachelard G. La formation de l'esprit scientifique. Librairie philosophique J.Vrin, 2000.
4. Lecourt D. Dictionnaire d'histoire et philosophie des sciences, 4^e réédition. Paris, 2006.
5. Bezzaoucha A. Epidémiologie et Biostatistiques, OPU, 1996.
6. Jenicek M, Cleroux R. Epidémiologie: principes, techniques, applications. Maloine, 1987.
7. Landrion G. La recherche clinique. Ed Masson, 1995.
8. Morton R.F, Hebel J.R. Epidémiologie et Biostatistiques. Doin éditeurs, 1983.
9. Rumeau Rouquette C, Blondel B, Kaminski M, Bréart G. Epidémiologie : méthodes et pratiques. Flammarion Médecine Sciences, 1994.
10. Bouyer J, Hemon D, Cordier S et al. Epidémiologie, principes et méthodes quantitatives. Ed Inserm, 1995, 305-330.
11. Czernichow P, Chaperon J, Le Coutour X. Epidémiologie, connaissances et pratiques. Ed Masson, 2001.
12. Dabis F, Drucker J, Morena A. Epidémiologie d'intervention. Ed Arnette, 1992.
13. Mayhall G. Hospital epidemiology and infection control. Ed Lippincott Williams and Wilkins, 2004.
14. Hennekens C.H, Buring J. Epidemiology in Medicine. Little Brown and Company, Boston, 1987.
15. Beaglehole R, Bonita R, Kjellstrom T. Basic Epidemiology. WHO, Geneva, 1993.
16. Greenberg R.S, Daniels S.R, Flanders W.D, Eley J.W, Boring J.R. Medical Epidemiology. Appleton and Lange, Norwalk, 1996.
17. Sackett D.L, Haynes, R.B, Guyatt G.H, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. Boston: Little, Brown and Company, 1991.

18. Von Wright GH. Explanation and understanding. London: Routledge and Kegan Paul, 1971.
19. Breuer J, Freud S. Studies in hysteria. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol 2. London: Hogarth Press, 1978.
20. Jaspers K. Allgemeine Psychopathologie. Berlin: Springer Verlag, 1913.
21. Baars B. The cognitive revolution in psychology. New York: Guildford, 1986.
22. Brown GW, Harris TO. Social origins of depression: a study of psychiatric disorder in women. London: Tavistock, 1978.
23. Main M, Hesse P. Lack of resolution of mourning in adulthood and its relationship to infant disorganisation: some speculations regarding causal mechanisms. Attachment in the preschool years. Chicago: University of Chicago Press; 1990, 161-82.
24. Brusset B. Diagnostic psychiatrique et différence du normal et du pathologique. EMC (Elsevier Masson SAS), Psychiatrie, 37-102-E-20, 1999.
25. Botbol M. Les modèles en psychiatrie. Le Livre blanc de la Fédération française de Psychiatrie. Editions John Libbey, 2003.
26. Mc Clellan JM, Werry JS. Research Psychiatric Diagnosis Interviews for Children and Adolescents. J Am Acad Child Adolesc Psychiatry, 2000.
27. Strauss J.S. Subjectivité et approche bio-psycho-sociale en psychiatrie. Annales Medico Psychologiques 2002, vol 160, 659-665.
28. Tordjman S. Évolution des recherches en pédopsychiatrie : vers la subjectivité et une approche intégrée psychodynamique et biologique appliquée à l'étude du développement normal et pathologique de l'enfant, Neuropsychiatrie de l'enfant, vol 54, 2006.
29. Prakash A. Qualitative research: an emergent paradigm for research in psychology. Psychol Stud 2001; 46 (3) :101-104.
30. Rennie DL, Watson KD, Monteiro AM. The rise of qualitative research in psychology. Can Psychol 2002; 43 (3) :179-189.
31. Early Clinical Drug Evaluation Unit. Assessment manual for psychopharmacology. National institute of mental health, 1976.
32. Rosenfeld J. Transformation, déformation, dysfonctionnement des interactions précoces. Psychiatr Enf 1990; 33 (2) :479-520.
33. De Roten Y, Fivaz E. Un guide pour l'évaluation du dialogue adulte-nourrisson. Psychiatr Enf 1992, 35(1) : 157-195.
34. Truffaut X. Méthodes, techniques et épistémologies de l'analyse quantitative. Pour la recherche. Bulletin de la fédération française de psychiatrie, Vol 42, Sept 2004.

35. Deloro C. La psychiatrie à l'ère des modèles. *L'Évolution Psychiatrique* 2010 ; 75(4) : 583-595.
36. Binet A, Simon T. Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *Année Psycho* 1905 ; 2 : 40-47.
37. Barthelemy C, Lelord G. Les échelles d'évaluation clinique en psychiatrie de l'enfant. Expansion Scientifique Française. Paris. 1991 ; 166p.
38. Dugas M. A propos des échelles et questionnaire d'évaluations en psychiatrie de l'enfant et de l'adolescent. *Neuropsychiatr Enfance Adolesc* 1986 ; 34 (2-3) : 127-129.
39. Dugas M, Le Heuzey M.F. Nosologies et systèmes de recueil des données en psychiatrie de l'enfant et de l'adolescent. *Psychiatrie de l'Enfant* 1988 ; 1 (XXXI) : 5-47.
40. Bouvard M, Cottraux J. Protocoles et échelles d'évaluation en psychiatrie et en psychologie. Ed. Masson, 2005 ; 327p.
41. Fleiss JL. Measuring agreement between two judges on the presence or absence of a trait. *Biometrics* 1975 ; 31 : 651-659.
42. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 ; 33 :159-174.
43. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ* 1992 ; 304 : 1491-1494.
44. Kramer MS, Feinstein AR. The biostatistics of concordance. *Clin Pharmacol Ther* 1981 ; 29 : 111-123.
45. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960 ; 20 : 27-46.
46. Fleiss J.L, Cohen J and Everitt B.S. Large sample standard errors of Kappa and weighted Kappa. *Psychol Bull* 1969 ; 72 : 323-327.
47. Fleiss J.L. Inference about weighted Kappa in the non-null case. *Appl Psychol Meas* 1978 ; 1 : 113-117.
48. Scott W.A. Reliability of content analysis : The case of nominal scale coding. *Public Opinion Q* 1955 ; 19 : 321-325.
49. Siegel S, Castellan N.J. Jr. *Non parametric Statistics for the Behavioral Sciences*. McGraw-Hill International Editions, 1988.
50. Watson P.F. Method agreement analysis: A review of correct methodology. *Theriogenology* 2010 ; 73 :1167–1179.
51. Gwet K. *Handbook of Inter-Rater Reliability*. Plastic Comb Editions, 2001, 180-185.

52. Sim J, Wright CC. The Kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005 ; 85(3) : 257-268.
53. Bergeri I, Michel R, Boutin J-P. Pour tout savoir ou presque sur le coefficient Kappa. *Med Trop* 2002 ; 62 : 634-636.
54. Santos F. Le Kappa de Cohen : un outil de mesure de l'accord inter-juges sur des caractères qualitatifs. CNRS, UMR 5199 ; sept 2010.
55. Obacz C. Clinique des phases prémorbides et prodromiques de la schizophrénie : étude rétrospective à propos de 50 cas. Thèse de Médecine 202 p. Université Henri Poincaré Nancy 1, Juillet 2010.
56. Brown C, Lloyd K. Qualitative methods in psychiatric research. *Advances in Psychiatric treatment*, 2001 ; 7 :350-356.
57. Hess U, Senécal S, Vallerand R.J. Les méthodes quantitative et qualitative de recherche en psychologie. *Méthodes de recherche en psychologie*. Boucherville, 2000.
58. Service des affections de longue durée et accords conventionnels, Haute Autorité de Santé, Guide ALD 23 « Schizophrénies », Juin 2007.
59. Hafner H, Riecher-Rossler A, Hambrecht, M, and al. IRAOS : an instrument for the assesment of onset and early course of schizophrenia. *Schizophr Res*. 1992 ; 6 : 209-223.
60. Schultze-Lutter F, Ruhrmann S, Berning J, and al. *Schizophr. Bull*. 2010 ; 36 : 182-191.
61. Holzer L, Halfon O, Jaugey L. Psychoses à l'adolescence, les neurosciences améliorent-elles la prédiction? *L'Encéphale*. 2008 ; 34 : 153-160.
62. Yung A.R, Yuen HP, Berger G, and al. Declining transition rate in ultra high risk (prodromal) services : dilution or reduction of risk? *Schizophr Bull*. 2007; 33 : 673-681.
63. Kendall M.G. Rank correlation methods. C. Griffin Booh Series, 1990.
64. Vanier M. Étude de fidélité d'un instrument de mesure de l'état de sujets traumatisés crâniens. Programme de traumatisme cranio-cérébral Centre Universitaire de santé McGill Hôpital Général de Montréal, Mars 2000.
65. Kottner J, Dassen T. Interpreting interrater reliability coefficients of the Braden scale: A discussion paper. Department of Nursing Science, Centre for Humanities and Health Sciences. Charite Universitats medizin Berlin, Germany, August 2007.
66. Byrt T, Bishop J, Carlin JB. Bias, prevalence and Kappa. *J Clin Epidemiol* 1993 ; 46 : 423-429.
67. Lantz C.A, Nebenzahl E. Behavior and intreprétation of the Kappa statistic: resolution of two paradoxes. *J Clin Epidemiol* 1996 ; 49 : 431-434.

68. Feinstein A.R, Cicchetti D.V. High Agreement but low Kappa: The problems of two paradoxes. *J Clin Epidem* 1990 ; 43 : 543-549.
69. Hoehler FK. Bias and prevalence effects on Kappa viewed in terms of sensitivity and specificity, *Journal of Clinical Epidemiology* 2000 ; 53 : 499-503.
70. Maclure M, Willett W. Misinterpretation and misuse of the Kappa statistic. *American Journal of Epidemiology* 1987 ; 126 : 161-169.
71. Viera A, Garrett J. Understanding Interobserver Agreement: The Kappa Statistic. *Fam Med* 2005 ; 37 : 360-363.

Annexe

Grille de cotation rétrospective dimensionnelle des phases précoces de la schizophrénie

Dimensions de l'enfance (de 0 à 12 ans)

Dimension cognitive/développementale :

Présence / Absence

- QI bas (< ou =à 69)..........
- Retard du développement moteur..........
- Retard de langage..........
- Troubles de la latéralisation..........
- Autres*

** troubles de la mémorisation, de la concentration, de l'attention, difficultés de raisonnement*

Dimension physique :

Présence / Absence

- Anorexie, perte de poids..........
- Troubles du sommeil
- Manifestations psychosomatiques..........

- Enurésie..........

Dimension négative :

Présence / Absence

- Retrait..........
- Difficultés relationnelles..........
- Stéréotypies motrices..........

Dimension positive/de discordance :

Présence / Absence

- Bizarreries comportementales..........
- Fabulations envahissantes..........

Dimension d'anxiété :

Présence / Absence

- Anxiété/ anxiété de séparation
- Eléments du registre obsessionnel..........
- Eléments du registre phobique..........

Dimension comportementale/d'impulsivité :

Présence / Absence

- Passages à l'acte auto-hétéro/agressifs..........
- Troubles du comportement non spécifiés..........

➤ Autres*

** impulsivité, intolérance à la frustration, opposition, irritabilité, agitation, colères*

Dimensions de l'adolescence (de 12 à 18ans)

Dimension cognitive :

Présence / Absence

➤ Troubles de l'attention, de la concentration.....

➤ Difficultés/diminution dans les capacités d'abstraction.....

➤ Lenteur idéative.....

➤ Autres*

** maladresse gestuelle, troubles des fonctions instrumentales*

Dimension physique :

Présence / Absence

➤ Anorexie, perte de poids non organique.....

➤ Troubles du sommeil.....

➤ Incurie.....

➤ Manifestations psychosomatiques.....

Dimension négative :

Présence / Absence

- Retrait.....
- Froideur affective.....
- Avolition
- Alogie.....

Dimension positive/de discordance :

Présence / Absence

- Discordance.....
- Hallucinations.....
- Persécution, suspicion
- Autres idées délirantes*.....

** de référence, mégalomaniaques, de culpabilité, d'empoisonnement...*

Dimension d'anxiété :

Présence / Absence

- Anxiété/ attaque de panique.....

- Éléments du registre obsessionnel..........
- Éléments du registre phobique..........

Dimension comportementale/d'impulsivité :

Présence / Absence

- Passages à l'acte auto-hétéro/agressifs..........
- Comportement délictueux..........
- Autres*..........

** intolérance à la frustration, fugues, troubles du comportement non spécifiés dans les dossiers (la mention « trouble du comportement » apparait dans le dossier sans être détaillée), agressivité sans passage à l'acte, opposition, agitation*

Dimension thymique :

Présence / Absence

- Items dépressifs..........
- Idées noires..........
- Fluctuations thymiques..........

RESUME DE LA THESE

Les manifestations prémorbides et prodromiques de la schizophrénie connaissent un regain d'intérêt de puis ces 15 dernières années. Une étude réalisée en 2008 au Centre Psychothérapeutique de Nancy (CPN) a investigué, à l'aide d'une grille de lecture, la présence de manifestations précoces de la maladie dans une population de 50 patients schizophrènes adultes.

Formalisée ensuite en grille de lecture complexe, notre travail a consisté à réutiliser cet instrument sur les mêmes 50 dossiers par un nouveau cotateur afin d'étudier la reproductibilité des données en mesurant la fidélité inter-juges pour les 48 items recueillis.

La première partie de ce travail rappelle le cadre théorique de la recherche en particulier en psychiatrie ainsi que les méthodes de construction et validation d'une échelle.

En second lieu, nous avons examiné la fidélité inter-juges du relevé de symptômes par le calcul et l'interprétation du coefficient Kappa de Cohen. Enfin, ces données ont abouti à envisager des modifications de l'instrument de mesure complexe utilisé dans cette étude de recherche.

TITRE EN ANGLAIS

Inter-rater reliability of a complex measuring instrument used in premorbid and prodromal phases of schizophrenia. A retrospective study of 50 patients.

THESE : MEDECINE SPECIALISEE-ANNEE 2012

MOTS CLEFS

Fidélité inter-juges, coefficient Kappa de Cohen, schizophrénie, prodromes, phases prémorbides.

ADRESSE DE L'UFR :

Université de Lorraine

Faculté de Médecine de Nancy

9, avenue de la forêt de Haye

54504 VANDOEUVRE-LES-NANCY Cedex