



Revue systématique de la littérature sur l'efficacité de l'amitriptyline dans les indications “ céphalées de tension ” et “ fibromyalgie ”

Valérie Guichardant, Houdna Laïb

► To cite this version:

Valérie Guichardant, Houdna Laïb. Revue systématique de la littérature sur l'efficacité de l'amitriptyline dans les indications “ céphalées de tension ” et “ fibromyalgie ”. Médecine humaine et pathologie. 2022. dumas-03710426

HAL Id: dumas-03710426

<https://dumas.ccsd.cnrs.fr/dumas-03710426>

Submitted on 30 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance.

La propriété intellectuelle du document reste entièrement celle du ou des auteurs. Les utilisateurs doivent respecter le droit d'auteur selon la législation en vigueur, et sont soumis aux règles habituelles du bon usage, comme pour les publications sur papier : respect des travaux originaux, citation, interdiction du pillage intellectuel, etc.

Il est mis à disposition de toute personne intéressée par l'intermédiaire de [l'archive ouverte DUMAS](#) (Dépôt Universitaire de Mémoires Après Soutenance).

Si vous désirez contacter son ou ses auteurs, nous vous invitons à consulter en ligne les annuaires de l'ordre des médecins, des pharmaciens et des sages-femmes.

Contact à la Bibliothèque universitaire de Médecine Pharmacie de Grenoble :

bump-theses@univ-grenoble-alpes.fr

Année : 2022

**REVUE SYSTÉMATIQUE DE LA LITTÉRATURE SUR L'EFFICACITÉ
DE L'AMITRIPTYLINE DANS LES INDICATIONS
« CÉPHALÉES DE TENSION » ET « FIBROMYALGIE »**

THÈSE
POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE
SPÉCIALITÉ : MÉDECINE GÉNÉRALE

Par Mme Valérie GUICHARDANT

[Données à caractère personnel]

Et Mme Houdna LAIB

[Données à caractère personnel]

THÈSE SOUTENUE PUBLIQUEMENT À LA FACULTÉ DE MÉDECINE DE GRENOBLE

Le 21/06/2022

DEVANT LE JURY COMPOSÉ DE :

Président du jury :

M. le Pr émérite Jacques DEMONGEOT

Membres :

M. le Pr Arnaud SEIGNEURIN

M. le Pr Guillaume ROYER DE VÉRICOURT

M. le Dr Yves BARO (directeur de thèse)

M. le Dr Matthieu REYMOND

L'UFR de Médecine de Grenoble n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions sont considérées comme propres à leurs auteurs.

CORPS	NOM-PRENOM	Discipline universitaire
PU-PH	ALBALADEJO Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH	ARVIEUX-BARTHELEMY Catherine	Chirurgie viscérale et digestive
PU-PH	BAILLET Athan	Rhumatologie
PU-PH	BARONE-ROCHETTE Gilles	Cardiologie
PU-PH	BAYAT Sam	Physiologie
MCF Ass.MG	BENDAMENE Farouk	Médecine Générale
PU-PH	BENHAMOU Pierre-Yves	Endocrinologie, diabète et maladies métaboliques
PU-PH	BERGER François	Biologie cellulaire
MCU-PH	BETRY Cécile	Nutrition
MCU-PH	BIDART-COUTTON Marie	Biologie cellulaire
PU-PH	BIOULAC-ROGIER Stéphanie	Pédopsychiatrie ; addictologie
PU-PH	BLAISE Sophie	Chirurgie vasculaire ; médecine vasculaire
PR Ass. Méd.	BOILLOT Bernard	Urologie
MCU-PH	BOISSET Sandrine	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	BONAZ Bruno	Gastroentérologie ; hépatologie ; addictologie
PU-PH	BONNETERRE Vincent	Médecine et santé au travail
PU-PH	BOREL Anne-Laure	Nutrition
PU-PH	BOSSON Jean-Luc	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	BOTTARI Serge	Biologie cellulaire
PR Ass.MG	BOUCHAUD Jacques	Médecine Générale
PU-PH	BOUGEROL Thierry	Psychiatrie d'adultes
PU-PH	BOUILLET Laurence	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
MCU-PH	BOUSSAT Bastien	Epidémiologie, économie de la santé et prévention
PU-PH	BOUZAT Pierre	Anesthésiologie-réanimation et médecine péri-opératoire
PU-PH émérite	BRAMBILLA Christian	Pneumologie
PU-PH émérite	BRAMBILLA Elisabeth	Anatomie et cytologie pathologiques
MCU-PH	BRENIER-PINCHART Marie Pierre	Parasitologie et mycologie
PU-PH	BRICAULT Ivan	Radiologie et imagerie médicale
PU-PH	BRICHON Pierre-Yves	Chirurgie thoracique et cardiovasculaire
MCU-PH	BRIOT Raphaël	Thérapeutique-médecine de la douleur ; Addictologie
PU-PH émérite	CAHN Jean-Yves	Hématologie
PU-PH émérite	CARPENTIER Patrick	Chirurgie vasculaire, médecine vasculaire
PR Ass.MG	CARRILLO Yannick	Médecine Générale
MCU-PH	CASPAR Yvan	Bactériologie-virologie ; hygiène hospitalière
PU-PH	CESBRON Jean-Yves	Immunologie
PU-PH	CHABARDES Stephan	Neurochirurgie
PU-PH	CHABRE Olivier	Endocrinologie, diabète et maladies métaboliques
PU-PH	CHAFFANJON Philippe	Anatomie
MCF Ass.MG	CHAMBOREDON Benoît	Médecine Générale
PU-PH	CHARLES Julie	Dermato-vénérologie
MCF Ass.MG	CHAUVET Marion	Médecine Générale

PU-PH	CHAVANON Olivier	Chirurgie thoracique et cardiovasculaire
PU-PH	CHIQUET Christophe	Ophthalmologie
PU-PH	CHIRICA Mircea	Chirurgie viscérale et digestive
PU-PH	CINQUIN Philippe	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	CLAVARINO Giovanna	Immunologie
MCU-PH	CLIN CHERPEC Rita	Nutrition
PU-PH	COHEN Olivier	Histologie, embryologie et cytogénétique
PU-PH	COSTENTIN Charlotte	Gastroentérologie ; hépatologie ; addictologie
PU-PH	COURVOISIER Aurélien	Chirurgie infantile
PU-PH	COUTTON Charles	Génétique
PU-PH	COUTURIER Pascal	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	CRACOWSKI Jean-Luc	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH	DEBATY Guillaume	Médecine d'Urgence
PU-PH	DEBILLON Thierry	Pédiatrie
PU-PH	DECAENS Thomas	Gastroentérologie ; hépatologie ; addictologie
PR Ass. Méd.	DEFAYE Pascal	Cardiologie
PU-PH	DEGANO Bruno	Pneumologie ; addictologie
PU-PH	DEMATTEIS Maurice	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
PU-PH émérite	DEMONGEOT Jacques	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	DERANSART Colin	Physiologie
PU-PH	DESCOTES Jean-Luc	Urologie
PU-PH	DETANTE Olivier	Neurologie
MCU-PH	DIETERICH Klaus	Génétique
MCU-PH	DOUTRELEAU Stéphane	Physiologie
PU-PH	EPAULARD Olivier	Maladies infectieuses ; Maladies tropicales
PU-PH	ESTEVE François	Biophysique et médecine nucléaire
MCU-PH	EYSSERIC Hélène	Médecine légale et droit de la santé
PU-PH émérite	FAGRET Daniel	Biophysique et médecine nucléaire
PU-PH	FAUCHERON Jean-Luc	Chirurgie viscérale et digestive
MCU-PH	FAURE Julien	Biochimie et biologie moléculaire
PU-PH	FERRETTI Gilbert	Radiologie et imagerie médicale
PU-PH	FONTAINE Éric	Nutrition
PU-PH	FRANCOIS Patrice	Epidémiologie, économie de la santé et prévention
MCU-MG	GABOREAU Yoann	Médecine Générale
PU-PH	GARBAN Frédéric	Hématologie ; Transfusion
PU-PH	GAUDIN Philippe	Rhumatologie
PU-PH	GAVAZZI Gaétan	Médecine interne ; gériatrie et biologie du vieillissement ; addictologie
PU-PH	GAY Emmanuel	Neurochirurgie
MCU-PH	GIAI Joris	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	GILLOIS Pierre	Biostatistiques, informatique médicale et technologies de communication
PU-PH	GIOT Jean-Philippe	Chirurgie plastique, reconstructrice et esthétique ; Brûlologie
MCU-PH	GRAND Sylvie	Radiologie et imagerie médicale
PU-PH émérite	GRIFFET Jacques	Chirurgie infantile
PU-PH	HAINAUT Pierre	Biochimie et biologie moléculaire
PU-PH émérite	HALIMI Serge	Nutrition
PU-PH	HENNEBICQ Sylviane	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PU-PH	HOFFMANN Pascale	Gynécologie-obstétrique ; gynécologie médicale

PU-PH émérite	HOMMEL Marc	Neurologie
PU-MG	IMBERT Patrick	Médecine Générale
PU-PH émérite	JOUK Pierre-Simon	Génétique
PU-PH	KAHANE Philippe	Physiologie
MCU-PH	KASTLER Adrian	Radiologie et imagerie médicale
PU-PH	KRAINIK Alexandre	Radiologie et imagerie médicale
PU-PH	LABARERE José	Epidémiologie, économie de la santé et prévention
MCU-PH	LABLANCHE Sandrine	Endocrinologie, diabète et maladies métaboliques
PU-PH	LANDELLE Caroline	Bactériologie – virologie ; Hygiène hospitalière
PU-PH	LANTUEJOUL Sylvie	Anatomie et cytologie pathologiques
PR Ass. Méd.	LARAMAS Mathieu	Cancérologie ; radiothérapie
MCU-PH	LARDY Bernard	Biochimie et biologie moléculaire
MCU-PH	LE GOUELLEC LE PISSART Audrey	Biochimie et biologie moléculaire
PU-PH	LECCIA Marie-Thérèse	Dermato-vénéréologie
PR Ass.MG	LEDOUX Jean-Nicolas	Médecine Générale
PU-PH émérite	LETOUBLON Christian	Chirurgie viscérale et digestive
PU-PH	LEVY Patrick	Physiologie
PU-PH	LONG Jean-Alexandre	Urologie
MCU-PH	LUPO Julien	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MAITRE Anne	Médecine et santé au travail
MCU-PH	MARLU Raphaël	Hématologie ; Transfusion
PR Ass. Méd.	MATHIEU Nicolas	Gastroentérologie ; hépatologie ; addictologie
MCU-PH	MAUBON Danièle	Parasitologie et mycologie
PU-PH	MAURIN Max	Bactériologie-virologie ; Hygiène hospitalière
MCU-PH	MC LEER Anne	Histologie, embryologie et cytogénétique
PR Ass. Méd	MICHY Thierry	Gynécologie-obstétrique
MCU-PH	MONDET Julie	Histologie, embryologie et cytogénétique
PU-PH	MORAND Patrice	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MOREAU-GAUDRY Alexandre	Biostatistiques, informatique médicale et technologies de communication
PU-PH	MORO Elena	Neurologie
PU-PH	MORO-SIBLOT Denis	Pneumologie ; addictologie
MCU-PH	MORTAMET Guillaume	Pédiatrie
PU-PH	MOUSSEAU Mireille	Cancérologie ; radiothérapie
PU-PH émérite	MOUTET François	Chirurgie plastique, reconstructrice et esthétique ; brûlologie
MCF Ass.MG	ODDOU Christel	Médecine Générale
PR Ass. Méd.	ORMEZANO Olivier	Cardiologie
MCU-PH	PACLET Marie-Hélène	Biochimie et biologie moléculaire
PU-PH	PAILHE Régis	Chirurgie orthopédique et traumatologie
PU-PH	PALOMBI Olivier	Anatomie
PU-PH	PARK Sophie	Hématologie ; Transfusion
PR Ass.MG	PAUMIER-DESBRIERES Françoise	Médecine Générale
PU-PH	PAYEN DE LA GARANDERIE Jean-François	Anesthésiologie-réanimation et médecine péri-opératoire
MCU-PH	PAYSANT François	Médecine légale et droit de la santé
MCU-PH	PELLETIER Laurent	Biologie cellulaire
PU-PH	PELLOUX Hervé	Parasitologie et mycologie
PU-PH	PEPIN Jean-Louis	Physiologie
PU-PH	PERARD DUMESTRE Chantal	Immunologie

PU-PH émérite	HOMMEL Marc	Neurologie
PU-MG	IMBERT Patrick	Médecine Générale
PU-PH émérite	JOUK Pierre-Simon	Génétique
PU-PH	KAHANE Philippe	Physiologie
MCU-PH	KASTLER Adrian	Radiologie et imagerie médicale
PU-PH	KRAINIK Alexandre	Radiologie et imagerie médicale
PU-PH	LABARERE José	Epidémiologie, économie de la santé et prévention
MCU-PH	LABLANCHE Sandrine	Endocrinologie, diabète et maladies métaboliques
PU-PH	LANDELLE Caroline	Bactériologie – virologie ; Hygiène hospitalière
PU-PH	LANTUEJOL Sylvie	Anatomie et cytologie pathologiques
PR Ass. Méd.	LARAMAS Mathieu	Cancérologie ; radiothérapie
MCU-PH	LARDY Bernard	Biochimie et biologie moléculaire
MCU-PH	LE GOUELLEC LE PISSART Audrey	Biochimie et biologie moléculaire
PU-PH	LECCIA Marie-Thérèse	Dermato-vénéréologie
PR Ass.MG	LEDoux Jean-Nicolas	Médecine Générale
PU-PH émérite	LETOUBLON Christian	Chirurgie viscérale et digestive
PU-PH	LEVY Patrick	Physiologie
PU-PH	LONG Jean-Alexandre	Urologie
MCU-PH	LUPO Julien	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MAITRE Anne	Médecine et santé au travail
MCU-PH	MARLU Raphaël	Hématologie ; Transfusion
PR Ass. Méd.	MATHIEU Nicolas	Gastroentérologie ; hépatologie ; addictologie
MCU-PH	MAUBON Danièle	Parasitologie et mycologie
PU-PH	MAURIN Max	Bactériologie-virologie ; Hygiène hospitalière
MCU-PH	MC LEER Anne	Histologie, embryologie et cytogénétique
PR Ass. Méd	MICHY Thierry	Gynécologie-obstétrique
MCU-PH	MONDET Julie	Histologie, embryologie et cytogénétique
PU-PH	MORAND Patrice	Bactériologie-virologie ; Hygiène hospitalière
PU-PH	MOREAU-GAUDRY Alexandre	Biostatistiques, informatique médicale et technologies de communication
PU-PH	MORO Elena	Neurologie
PU-PH	MORO-SIBILOT Denis	Pneumologie ; addictologie
MCU-PH	MORTAMET Guillaume	Pédiatrie
PU-PH	MOUSSEAU Mireille	Cancérologie ; radiothérapie
PU-PH émérite	MOUTET François	Chirurgie plastique, reconstructrice et esthétique ; brûlologie
MCF Ass.MG	ODDOU Christel	Médecine Générale
PR Ass. Méd.	ORMEZANO Olivier	Cardiologie
MCU-PH	PACLET Marie-Hélène	Biochimie et biologie moléculaire
PU-PH	PAILHE Régis	Chirurgie orthopédique et traumatologie
PU-PH	PALOMBI Olivier	Anatomie
PU-PH	PARK Sophie	Hématologie ; Transfusion
PR Ass.MG	PAUMIER-DESBRIERES Françoise	Médecine Générale
PU-PH	PAYEN DE LA GARANDERIE Jean-François	Anesthésiologie-réanimation et médecine péri-opératoire
MCU-PH	PAYSANT François	Médecine légale et droit de la santé
MCU-PH	PELLETIER Laurent	Biologie cellulaire
PU-PH	PELLOUX Hervé	Parasitologie et mycologie
PU-PH	PEPIN Jean-Louis	Physiologie
PU-PH	PERARD DUMESTRE Chantal	Immunologie

PU-PH	PERENNOU Dominique	Médecine physique et de réadaptation
PU-PH	PERNOD Gilles	Chirurgie vasculaire ; Médecine vasculaire
MCF	PINSAULT Nicolas	Sciences de la rééducation et de réadaptation
PU-PH	PIOLAT Christian	Chirurgie infantile
PU-PH	PISON Christophe	Pneumologie ; Addictologie
PU-PH	PLANTAZ Dominique	Pédiatrie
PU-PH	POIGNARD Pascal	Bactériologie-virologie ; Hygiène hospitalière
PU-PH émérite	POLACK Benoît	Hématologie ; Transfusion
PU-PH	POLOSAN Mircea	Psychiatrie d'adultes ; Addictologie
PU-PH émérite	RAMBEAUD Jean-Jacques	Urologie
PU-PH	RAY Pierre	Biologie et médecine du développement et de la reproduction ; gynécologie médicale
PR Ass. Méd.	RECHE Fabian	Chirurgie viscérale et digestive
MCU-PH	RENDU John	Biochimie et biologie moléculaire
MCU-PH émérite	RIALLE Vincent	Biostatistiques, informatique médicale et technologies de communication
PU-PH	RIETHMULLER Didier	Gynécologie-obstétrique ; gynécologie médicale
PU-PH	RIGHINI Christian	Oto-rhino-laryngologie
PU-PH émérite	ROMANET Jean Paul	Ophtalmologie
PU-PH	ROSTAING Lionel	Néphrologie
PU-PH	ROUSTIT Matthieu	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
MCU-PH	ROUX-BUISSON Nathalie	Biochimie et biologie moléculaire
PR Ass.MG	ROYER DE VERICOURT Guillaume	Médecine Générale
PU-PH émérite	SARAGAGLIA Dominique	Chirurgie orthopédique et traumatologie
MCU-PH	SATRE Véronique	Génétique
PU-PH	SAUDOU Frédéric	Biologie cellulaire
PU-PH	SCHMERBER Sébastien	Oto-rhino-laryngologie
PU-PH	SCHWEBEL Carole	Médecine intensive-réanimation
PU-PH	SCOLAN Virginie	Médecine légale et droit de la santé
PU-PH	SEIGNEURIN Arnaud	Epidémiologie, économie de la santé et prévention
PU-PH	SPEAR Rafaëlle	Chirurgie vasculaire ; médecine vasculaire
PU-PH émérite	STAHL Jean-Paul	Maladies infectieuses ; Maladies tropicales
PU-PH	STANKE Françoise	Pharmacologie fondamentale ; pharmacologie clinique ; addictologie
MCU-PH	STASIA Marie-José	Biochimie et biologie moléculaire
PU-PH	STURM Nathalie	Anatomie et cytologie pathologiques
PU-PH	TAMISIER Renaud	Physiologie
PU-PH	TERZI Nicolas	Médecine intensive-réanimation
PU-PH	THEVENON Julien	Génétique
PU-PH	TOFFART Anne-Claire	Pneumologie ; Addictologie
PU-PH	TONETTI Jérôme	Chirurgie orthopédique et traumatologie
PU-PH	TOUSSAINT Bertrand	Biochimie et biologie moléculaire
PU-PH	VALMARY-DEGANO Séverine	Anatomie et cytologie pathologiques
PU-PH	VANZETTO Gérald	Cardiologie
PU-PH	VUILLEZ Jean-Philippe	Biophysique et médecine nucléaire
PU-PH	WEIL Georges	Epidémiologie, économie de la santé et prévention
PU-PH	ZAOUI Philippe	Néphrologie
PU-PH émérite	ZARSKI Jean-Pierre	Gastroentérologie ; hépatologie ; addictologie

Remerciements:

A notre président de jury, Monsieur le Professeur Jacques DEMONGEOT : nous sommes très honorées de l'intérêt que vous portez à notre travail et du temps que vous avez pris pour nous rencontrer. Nous tenons à vous exprimer toute notre reconnaissance pour l'aide apportée dans la constitution de notre jury de thèse. Nous vous remercions avec gratitude d'avoir accepté de présider ce jury.

A Monsieur le Professeur Arnaud SEIGNEURIN : vous avez reçu notre demande assez tardivement et cependant vous avez su vous rendre disponible. Nous vous en remercions sincèrement. Merci de l'honneur que vous nous faites en acceptant de participer à notre jury de thèse.

A Monsieur le Professeur Guillaume ROYER de VERICOURT : merci de l'honneur que vous nous avez fait en acceptant d'abord d'assister à la présentation de notre mémoire, puis de participer à notre jury de thèse.

A Monsieur le Docteur Matthieu REYMOND : merci de votre disponibilité et de votre bienveillance à l'égard de notre travail. Merci d'avoir accepté de faire partie de notre jury et d'apporter une critique à notre travail.

A notre directeur de thèse, et à mon maître de stage Monsieur le Docteur Yves BARO : nous connaissons l'importance qu'à dans ta pratique médicale l'Evidence Based Medecine. C'est donc naturellement à toi que nous avons proposé d'encadrer notre travail. Nous te remercions d'avoir accepté, ainsi que pour ta disponibilité, tes conseils et ta bienveillance.

A Monsieur le Professeur Rémi BOUSSAGEON : merci pour vos conseils et vos explications concernant la méthode REB.

A Mme Mélanie VELOSO biostatisticienne du pôle recherche du Centre Hospitalier Universitaire Grenoble Alpes : merci de nous avoir aidées dans la partie statistique de notre travail par votre disponibilité, vos explications claires, et votre réactivité.

A Madame le Docteur Adeline LINDE BOUCHET : merci pour tous tes conseils et ta disponibilité au cours de notre travail. Tu es une amie précieuse, toujours présente, que ce soit dans les meilleurs moments comme lors de nos voyages, ou dans les difficultés.

A Madame le Docteur Theresa DAILY : merci pour ton aide en particulier linguistique et de ton soutien pendant notre travail. Ton calme, ta soif de nouvelles expériences, ta générosité et ton souci du bien-être d'autrui font de toi une amie extraordinaire.

A tous nos maîtres de stage, qui se sont investis pour nous transmettre leurs connaissances et leur passion de la médecine générale, pour nous permettre de devenir les médecins que nous sommes maintenant. A tous les services où nous sommes passées en stage et qui ont contribué à notre formation.

Remerciements de Valérie GUICHARDANT :

A mes parents : vous avez été acteurs de ma réussite que ce soit par votre soutien logistique comme les aller-retours au Chartreux, ou par votre investissement personnel tout au long de mon parcours avec en dernier lieu le temps que vous avez consacré à la relecture de notre thèse. Vous avez veillé à ce que je ne manque de rien et à ce que je bénéficie des meilleures conditions possibles pour réussir. Vous m'avez transmis vos goûts pour les voyages, la randonnée, la nature, et la bonne cuisine ainsi que vos valeurs, en particulier pour le travail bien fait avec cette devise : « le travail paye toujours ». Merci pour votre amour.

A ma grand-mère « Mamie Malafretaz », tu as toujours été présente et discrète. Merci pour ton soutien et tes prières.

A ma grand-mère « Mamie Belley » : tu es partie depuis quelques temps déjà, mais je voulais te remercier pour ta gentillesse et ta générosité.

A mon grand-père : je te remercie de tous les bons moments passés ensemble, les après-midis au lac où tu nous emmenais lorsque nous étions enfants, les restaurants et les randonnées... Tu m'as appris qu'il y avait toujours des solutions, quelque chose à faire ou à bricoler pour résoudre ou contourner le problème ou la difficulté.

A Romain BAUD : c'est dans les plaisirs du quotidien et les attentions l'un envers l'autre que se trouve le bonheur, et notre amour. C'est ainsi que ma vie se déroule à tes côtés. Merci pour tous les super bons moments passés ensemble, et pour la découverte d'une autre façon de randonner et de voyager grâce au VTT.

A Maryse et Jean-Louis BAUD : merci pour votre accueil, votre gentillesse et votre soutien.

A Madame le Docteur Anne BURDIN : je te remercie d'avoir cru en moi, de m'avoir encouragée et donné les moyens de devenir médecin. Ta pratique médicale alliant compétence et humanité est un modèle pour moi.

A Madame le Docteur Christel ODDOU : je vous remercie pour votre soutien, votre bienveillance et votre implication dans mon parcours. Vous avez été une personne ressource essentielle durant mon internat.

A Monsieur le Professeur Patrick IMBERT : je vous remercie pour votre préoccupation du bien-être de vos étudiants.

A Madame le Docteur Isabelle CHIRPAZ-CERBAT : je vous remercie pour votre écoute et votre soutien.

A Madame le Docteur Christelle HUET : merci pour ton aide et ton soutien pendant l'externat et en particulier pendant la P1. Tu as été une marraine formidable.

A tous mes amis des Chartreux puis de l'externat puis de l'internat et enfin d'Annecy : Roxanne, Edouard, Sixtine, Luc, Dimitri, Victor, Thanh Danh, Clotilde et Mathieu, David et Anastasia, Federico et Fedora, Julien, Scrat et Lisa. Vous avez en commun d'être plein de bienveillance et d'humanité. Je vous remercie pour tous les moments passés ensemble.

A Marine THOMAS : tu es une amie depuis plus de 20 ans, depuis la 6^{ème} précisément. Nous avons grandi ensemble, partagé le même univers fantastique des livres, le même collège et le même lycée. Le temps et nos vies n'ont fait que nous rapprocher. Merci pour ton amitié

A Sixtine VAN HILLE : rencontrée aussi en 6^{ème} sur les bancs des Chartreux notre passion pour la lecture, notre goût pour les voyages et la montagne, notre foi et nos métiers ont créé le socle de notre amitié. Comme le bon vin ou le bon cognac, le temps la bonifie.

A Dr Etienne MASSARDIER : mon spécialiste préféré, ma hotline traumato, tu es un ami génial, fiable, organisé, consciencieux et calme dans ton travail comme dans les moments entre amis. Merci pour tes conseils et les bons moments passés ensemble

A Dr Valentin MASSARDIER : merci pour ton dynamisme, ta bonne humeur, tes encouragements dans les moments difficiles que ce soit en randonnée ou dans la vie. Merci pour ton amitié

A Dr Tommy STUDER : ton symbole c'est un smiley qui sourit. Un rien tu le transformes en quelque chose de drôle. Tu m'as fait découvrir des coins de paradis dans les montagnes. Tu es aussi une personne très humaine et compétente et tu m'as beaucoup aidée. Merci pour ta bonne humeur et ta bienveillance.

A Valentine DURIVAL : l'internat nous a permis de nous découvrir une passion commune pour la randonnée et le trek. Tu apportes aux projets auxquels tu participes une organisation et un réalisme qui leurs permet d'aboutir. Merci pour ton amitié

A Mélissa BOUDIA : merci les beaux moments passés ensemble. J'aime beaucoup ta joie de vivre, ta volonté, ton épanouissement dans ta vie professionnelle, la diversité de tes centres d'intérêt, tes efforts pour rassembler tes amis et nous faire découvrir des lieux magnifiques.

A Lucie BONNICHON : merci pour toutes les sorties, week-end, vacances que tu as largement contribué à organiser. Tu es une amie sur laquelle je peux compter et une super « coloc » : drôle, gentille, et agréable à vivre.

A Dr Anne LEBEUGLE : notre amitié est née lors de notre rencontre au premier jour de l'internat. Tu es celle à qui arrivent des mésaventures incroyables que tu sais si bien raconter. Tu es une personne pleine de ressources, battante et une super amie.

A ma co-thésarde Houdna LAIB : travailler ensemble m'a permis de découvrir une personne formidable, pleine de bienveillance, et doter d'une patience à toute épreuve. Tu es la personne parfaite avec qui collaborer : jamais dans le jugement, toujours positive. Ton mari est aussi quelqu'un de très gentil et soucieux du bien-être des gens, et je suis heureuse de l'avoir rencontré. Votre accueil et votre générosité caractérisent votre famille. Merci

Remerciements de Houdna LAIB:

A ma mère, je ne serais pas la personne que je suis aujourd'hui sans toi. Tu as toujours été présente pour moi, tu as toujours fait en sorte que je ne manque de rien et je te remercie pour cela. Ma réussite est également la tienne.

A mon père, je te remercie pour tous tes encouragements tout au long de mon parcours et pour tous les moments passés avec toi.

A mon mari, mon meilleur ami, mon confident, l'amour de ma vie, le père de mes enfants. Merci de me supporter au quotidien et d'être là pour moi dans les bons et les mauvais moments.

A mes petits trésors, Maria Nour et Adam Abdelaziz, je vous remercie pour tout le bonheur que vous avez apporté dans ma vie. Je suis tellement heureuse d'être votre maman et d'avoir la chance de vous voir grandir jour après jour.

A Blanche, je vous remercie pour votre bienveillance à l'égard de ma fille, sans vous ce travail n'aurait pas été possible.

A ma tante Louisa et à mon oncle Mabrouk, je vous remercie pour toute l'affection que vous m'avez apportée.

A mes frères et sœurs de cœur, Farouk, Fouzia, Salim et Rahim, merci d'être toujours présents pour moi et cela malgré la distance. A Nada et à Lilou "le p'tit sucre", j'ai tellement hâte de te voir grandir et jouer avec Maria et Adam.

A mes oncles et tantes maternelles, mes cousins et cousines, je vous remercie tous pour tous ces étés passés ensemble et ces beaux souvenirs à vos côtés.

A ma belle-mère, je te remercie de me considérer comme ta propre fille et d'être aussi bienveillante envers moi.

A mon beau-père, j'aurais tellement aimé que tu sois présent à nos côtés aujourd'hui.

A mes beaux-frères, je vous remercie de m'avoir aussi bien accueillie dans votre famille et de me considérer comme votre sœur. A oncle L'Aïd, je te remercie pour ta gentillesse et ta bonne humeur.

A toutes mes amies, à Asmae, Mariam, Lydia, Nadia, Selma je vous remercie pour tous ces bons moments passés à vos côtés. A tous mes camarades d'école, de collège, de lycée, de prépa, erasmus, à mes co-externes, mes co-internes, je vous remercie pour tous ces beaux souvenirs ensemble.

A ma co-thésarde : J'ai été heureuse de travailler sur cette thèse avec toi. Au début nous nous connaissions à peine et au fil du temps une belle amitié est née. Merci d'être la personne que tu es et surtout ne changes jamais !

Le serment d'Hippocrate

Texte revu par l'Ordre des médecins en 2012

“ **Au moment d'être admis(e) 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(e) dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu(e) 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(e) et méprisé(e) si j'y manque.

”



UNIVERSITÉ GRENOBLE ALPES
UFR DE MÉDECINE DE GRENOBLE

Année : 2022

Revue systématique de la littérature sur l'efficacité de l'amitriptyline dans les indications « céphalées de tension » et « fibromyalgie ».

Systematic review of the literature on the effectiveness of amitriptyline in the indications « tension headache » and « fibromyalgia ».

GUICHARDANT Valérie et LAIB Houdna

MOTS CLÉS MESH : Amitriptyline - Céphalée de tension chronique - Fibromyalgia

MESH TERMS / KEYWORDS : Amitriptyline – Chronic tension headache –Fibromyalgia

FILIÈRE : Médecine Générale.

Les auteurs déclarent n'avoir aucun conflit d'intérêt.

Ce travail de thèse n'a reçu aucun financement ni compensation.

RÉSUMÉ :

Contexte : L'amitriptyline est recommandée dans le traitement de céphalées de tension chroniques en première intention et utilisée hors AMM pour la fibromyalgie. L'objectif de cette étude est de réaliser une revue systématique de la littérature avec une méta-analyse des essais cliniques randomisés comparant l'efficacité de l'amitriptyline à un placebo dans le traitement des céphalées de tension et de la fibromyalgie chez l'adulte.

Méthode : Une recherche bibliographique a été réalisée sur les bases Medline, ClinicalTrials.gov et CENTRAL. Seuls les essais randomisés évaluant l'amitriptyline contre placebo, en double aveugle en français ou en anglais ont été inclus. Le risque de biais a été évalué pour chaque étude à l'aide de l'outil Cochrane Risk of bias 2 (Rob2). Le niveau de preuve a été évalué selon la méthode REB et selon GRADE.

Résultats : Pour l'indication céphalées de tension 7 études ont été incluses. Toutes étaient à risque élevé de biais. Seule une analyse qualitative a été possible. Celle-ci suggère une différence en faveur de l'amitriptyline. Pour l'indication fibromyalgie, 13 études ont été incluses, 1 étude a été exclue secondairement. Toutes étaient à risque élevé de biais. L'analyse quantitative retrouve une différence statistiquement significative en faveur de l'amitriptyline pour l'évaluation du ressenti global du patient, l'amélioration global selon le médecin, l'intensité de la douleur et le sommeil. Les résultats suggèrent une différence en faveur de l'amitriptyline mais ne sont pas significatifs pour l'amélioration de la qualité de vie et la fatigue. L'évaluation du niveau de preuve pour les deux indications selon la méthode REB montre qu'il existe "un manque de preuve" et selon GRADE le niveau de preuve est "très faible".

Discussion : Cette revue de la littérature met en évidence un manque de preuve scientifique concernant l'efficacité de l'amitriptyline dans le traitement des céphalées de tension chroniques et de la fibromyalgie. Cela s'explique par la présence de risques de biais majeurs qui s'expliquent par le fait que les études soient anciennes, de faible puissance, avec de forts taux de données manquantes et la présence d'effets indésirables spécifiques rendant difficile le maintien du double aveugle. Nous préconisons donc la réalisation de nouvelles études de meilleure qualité méthodologique.

ABSTRACT :

Background: Amitriptyline is a recommended first-line treatment for chronic tension headaches and is also used for fibromyalgia, despite not being officially approved in this indication. This study aims to systematically review the literature and perform a meta-analysis of randomized clinical trials comparing the effectiveness of amitriptyline versus placebo in chronic tension headaches and fibromyalgia in adults.

Method: Medline, ClinicalTrials.gov and CENTRAL databases were searched for double-blind randomized trials comparing amitriptyline to placebo, written in English or French. Risk of bias was evaluated for each study using the Cochrane Risk of Bias 2 (Rob2) tool. Level of evidence was assessed using the REB method and the GRADE system.

Results: For chronic tension headaches, 7 studies were included, and all had a high risk of bias. Only a qualitative analysis could be performed, and it suggested a difference in favour of amitriptyline. For fibromyalgia, 13 studies were initially included, and 1 was later excluded. All had a high risk of bias. The quantitative analysis showed a statistically-significant difference in favour of amitriptyline for overall evaluation by patient, overall evaluation by attending physician, pain levels and sleep. The results also suggested an improvement under amitriptyline for quality of life and fatigue, but were not statistically significant. The evidence level was assessed by the REB method as “lack of evidence” and by the GRADE system as “very low”.

Discussion: This review of the literature shows a lack of scientific evidence for the effectiveness of amitriptyline in chronic tension headaches and fibromyalgia, mainly due to the high risk of bias: the studies are old, under-powered, have significant amounts of missing data and struggle to ensure double-blinding because of characteristic side effects. New studies with better methodology are therefore essential.

TABLE DES MATIÈRES

Céphalées de tension

I- INTRODUCTION	24
1. Contexte	24
2. Objectifs	25
II- MATERIELS ET METHODES	26
1. Protocole et enregistrement.....	26
2. Critères d'éligibilité	26
3. Sources d'information.....	26
4. Recherche.....	27
5. Sélection des études	27
6. Extraction des données	28
7. Risque de biais inhérent à chacune des études.....	28
8. Quantification des résultats	32
9. Synthèse des résultats	32
10. Risque de biais transversal aux études et évaluation du niveau de preuve	33
III- SELECTION DES ETUDES	34
IV- RÉSULTATS	36
1. Caractéristiques des études sélectionnées	36
2. Risque de biais relatif aux études.....	41
a) Les biais dus au processus de randomisation	42
b) Les biais dus à l'effet de période et à l'effet carry-over.....	43
c) Les biais dus aux écarts aux interventions prévues.....	43
d) Biais résultant des données manquantes	44
e) Biais résultant de la mesure du critère de jugement.....	45
f) Biais résultant de la sélection des résultats rapportés.....	45
g) Risque lié à la multiplicité des tests statistiques.....	45
3. Analyse de sensibilité	46
4. Résultats de chaque étude	49
a) Lance 1964	49
b) Göbel 1994	51

c) Pfaffenrath 1994.....	51
d) Brendtsen 1996.....	53
e) Diamond 1998.....	54
f) Brendtsen 2000.....	55
g) Ashina et Brendtsen 2004	56
5. Synthèse des résultats.....	57
a) Synthèse qualitative.....	57
b) Synthèse quantitative.....	61
6. Evaluation du niveau de preuve global selon REB et GRADE	62
a) Niveau de preuve pour “ la durée des céphalées “	63
b) Niveau de preuve pour “ l’intensité des céphalées “	64
c) Niveau de preuve pour “ la fréquence des céphalées “	65
V- DISCUSSION.....	66
1. Interprétation des résultats	66
2. Confrontation à la littérature	68
3. Limites	70
4. Implications pour la pratique	71
5. Implications pour la recherche.....	71
VI- CONCLUSION.....	73
VII- FINANCEMENT ET DÉCLARATION DE CONFLITS D'INTÉRÊTS	74
VIII- BIBLIOGRAPHIE	75
IX- ANNEXES.....	78
1. J.W Lance 1964: Treatment of chronic tension headache	78
a) Caractéristiques des études.....	78
b) Tableau RoB2.....	79
c) Résumé des domaines et jugement global selon RoB.....	84
d) Tableau de résultats.....	85
2. Gobel 1994: Chronic tension-type headache: amitriptyline reduces clinical headache duration and experimental pain sensitivity but does not alter pericranial muscle activity readings	87
a) Caractéristiques de l’étude	87
b) Tableau RoB 2.....	88
c) Résumé des domaines et jugement global.....	92
d) Tableau de résultats	93
3. Pfaffenrath 1994 : Efficacy and tolerability of amitriptylinexide in the treatment of chronic tension-type headache: a multi-centre controlled study	95

a) Caractéristiques de l'étude	95
b) Tableau RoB2.....	96
c) Résumé des domaines et jugement global.....	100
d) Tableau de résultats	101
4. Brendtsen 1996: A non-selective (amitriptyline), but not a selective citalopram), serotonin reuptake inhibitor is effective in the prophylactic treatment of chronic tension-type headache	104
a) Caractéristiques de l'étude	104
b) Tableau RoB 2.....	105
c) Résumé des domaines et jugement global.....	109
d) Tableau de résultats	110
5. Diamond 1998 : Chronic tension headache--treated with amitriptyline a double-blind study	112
a) Caractéristiques de l'étude	112
b) Tableau RoB2.....	113
c) Résumé des domaines et jugement global.....	117
d) Tableau de résultats	118
6. Brendtsen 2000: Amitriptyline reduces myofascial tenderness in patients with chronic tension-type headache	120
a) Caractéristiques de l'étude	120
b) Tableau RoB2.....	121
c) Résumé des domaines et jugement global.....	126
d) Tableau de résultats	127
7. ASHINA et BRENDTSEN 2004 : Analgesic effect of amitriptyline in chronic tension-type headache is not directly related to serotonin reuptake inhibition.	128
a) Caractéristiques de l'étude	128
b) Tableau RoB2.....	130
c) Résumé des domaines et jugement global.....	134
d) Tableau de résultats	135

Fibromyalgie

I-	INTRODUCTION	138
1.	Contexte	138
2.	Objectif	139
II-	MATERIELS ET METHODE.....	140
1.	Protocole et enregistrement.....	140
2.	Critères d'éligibilité	140
3.	Sources d'information.....	140
4.	Recherche.....	141
5.	Sélection des études	141
6.	Extraction des données	142
7.	Risque de biais inhérent à chacune des études.....	142
8.	Quantification des résultats.....	146
9.	Synthèse des résultats.....	146
10.	Risque de biais transversal aux études et évaluation du niveau de preuve	147
III-	SELECTION DES ETUDES	148
IV-	RESULTATS	151
1.	Caractéristiques des études sélectionnées	151
a)	Schéma d'étude	151
b)	Durée des études.....	152
c)	Caractéristiques des populations étudiées	152
d)	Analyse des résultats	154
e)	Critères de jugement.....	155
2.	Risque de biais	158
a)	Les biais dans le processus de randomisation	159
b)	Les biais dus à l'effet de période et à l'effet carry-over.....	161
c)	Les biais dus aux écarts aux interventions prévues.....	161
d)	Biais résultant des données manquantes	162
e)	Biais résultant de la mesure des critères de jugement	163
f)	Biais résultant de la sélection des résultats rapportés.....	163
g)	Risque lié à la multiplicité des tests statistiques.....	164
3.	Analyse de sensibilité:	166
4.	Résultats de chaque étude	173
a)	Carette 1986	174

b)	Don Goldenberg 1986	176
c)	Scudds 1989	177
d)	Carette 1994	178
e)	Kempenaers 1994.....	180
f)	Carette 1995	181
g)	Don Goldenberg 1996	182
h)	Ginsberg 1996	183
i)	Hannonen 1998	185
j)	Heymann 2001	186
k)	Braz 2012	188
5.	Synthèse des résultats	190
a)	Evaluation de l'amélioration globale des symptômes par le patient	193
a.1-	Le ressenti global du patient	193
a.2-	Les répercussions de la pathologie sur la qualité de vie du patient	195
b)	Evaluation de l'amélioration globale par le médecin	197
c)	Evaluation de la douleur.....	199
d)	Evaluation du sommeil.....	202
e)	Evaluation de la fatigue	203
f)	Evaluation des effets secondaires et des perdus de vue	206
a.1-	Les effets indésirables.....	206
a.2-	Les perdus de vue	207
6.	Evaluation du niveau de preuve global selon REB et GRADE	209
a)	Evaluation de l'amélioration globale des symptômes par le patient	210
a.1-	Ressenti du patient	210
a.2-	Répercussion de la pathologie sur la qualité de vie	212
b)	Evaluation de l'amélioration globale par le médecin.....	214
c)	Evaluation de la douleur.....	216
d)	Evaluation du Sommeil	218
e)	Evaluation de la fatigue.....	220
V-	DISCUSSION.....	223
1.	Interprétation des résultats	223
a)	Critère de jugement « amélioration globale évaluée par le patient »	225
b)	Critère de jugement « amélioration globale évaluée par le médecin »	227
c)	Critère de jugement « intensité de la douleur »	228
d)	Critère de jugement « troubles du sommeil ».....	229

e) Critère de jugement « fatigue »	230
f) Pour les effets indésirables et les perdus de vue	231
2. Confrontation à la littérature	233
3. Les limites de notre étude	237
4. Implication pour la pratique	238
5. Implication pour la recherche	239
VI- CONCLUSION.....	242
VII- FINANCEMENT ET DÉCLARATION DE CONFLITS D'INTÉRÊTS	243
VIII- BIBLIOGRAPHIE	243
IX- ANNEXES.....	246
1. S. Carette 1986: Evaluation of amitriptyline in primary fibrositis. A double-blind, placebo-controlled study	246
a) Caractéristiques de l'étude	246
b) Tableau RoB2.....	248
c) Résumé des domaines et jugement global.....	252
d) Tableaux de résultats	253
2. Don. L. Goldenberg 1986: A randomized, controlled trial of amitriptyline and naproxen in the treatment of patients with fibromyalgia	254
a) Caractéristiques de l'étude	254
b) Tableau RoB2.....	256
c) Résumé des domaines et jugement global.....	260
d) Tableau de résultats	261
3. Ra. Scudds 1989: Improvements in pain responsiveness in patients with fibrositis after successful treatment with amitriptyline.....	262
a) Caractéristiques de l'étude	262
b) Tableau RoB2.....	263
c) Résumé des domaines et jugement global.....	267
d) Tableau de résultats	268
4. R. Jaeschke 1991: Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials	269
a) Caractéristiques de l'étude	269
b) Tableau RoB2.....	270
c) Résumé des domaines et jugement global.....	274
d) Tableau de résultats	275
5. S. Carette 1994 : Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. A randomized, double-blind clinical trial	276

a) Caractéristiques de l'étude	276
b) Tableau RoB 2.....	278
c) Résumé des domaines et jugement global.....	282
d) Tableau de résultats	283
6. C. Kempenaers 1994: Effect of an antidiacephalon immune serum on pain and sleep in primary fibromyalgia	285
a) Caractéristiques de l'étude	285
b) Tableau RoB2.....	286
c) Résumé des domaines et jugement global.....	289
d) Tableau de résultats	290
7. S. Carette 1995: Sleep electroencephalography and the clinical response to amitriptyline in patients with fibromyalgia.....	293
a) Caractéristiques de l'étude	293
b) Tableau RoB2.....	294
c) Résumé des domaines et jugement global.....	298
d) Tableau de résultats	299
8. Don. L. Goldenberg 1996: A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia	300
a) Caractéristiques de l'étude	300
b) Tableau RoB2.....	302
c) Résumé des domaines et jugement global.....	306
d) Tableau de résultats	307
9. F. Ginsberg 1996 : A randomized placebo-controlled trial of sustained-release amitriptyline in primary fibromyalgia	308
a) Caractéristiques de l'étude	308
b) Tableau RoB2.....	310
c) Résumé des domaines et jugement global.....	313
d) Tableau de résultats	314
10. P. Hanonnen 1998: A randomized double-blind placebo controlled study of moclobemide and amitriptyline in fibromyalgia	319
a) Caractéristiques de l'étude	319
b) Tableau RoB2.....	321
c) Résumé des domaines et jugement global.....	325
d) Tableau des résultats	326
11. R.E Heymann 2001: A double-blind, randomized, controlled study of amitriptyline, nortriptyline and placebo in patients with fibromyalgia. An analysis of outcome measures.	

a) Caractéristiques de l'étude	328
b) Tableau RoB2.....	329
c) Résumé des domaines et jugement global.....	332
d) Tableau de résultats	333
12. E.A Fors 2002: The effect of guided imagery and amitriptyline on daily fibromyalgia pain: a prospective, randomized, controlled trial.	334
a) Caractéristiques de l'étude	334
b) Tableau RoB2.....	335
c) Résumé des domaines et jugement global.....	339
d) Tableau de résultats	340
13. A. Braz 2012: Effects of Panax ginseng extract in patients with fibromyalgia: a 12-week, randomized, double-blind, placebo-controlled trial.	342
a) Caractéristiques de l'étude	342
b) Tableau RoB 2.....	344
c) Résumé des domaines et jugement global.....	348
d) Tableau de résultats	349

Conclusion générale.....	350
--------------------------	-----

Erratum	355
---------------	-----

PARTIE 1 : CEPHALEES DE TENSION

PARTIE 1 : Céphalées de tension

I- INTRODUCTION

1. Contexte

Les céphalées de tensions concernent, selon les études, 30 à 78% des individus au cours de leur vie. Le plus souvent épisodiques (moins de 1 jour par mois), elles sont dites fréquentes (1 à 14 jours/mois) dans environ 30% des cas, et sont chroniques ou permanentes (plus de 15 jours par mois) dans 2 à 5 % des cas. (1)

Si les céphalées épisodiques n'aboutissent qu'à une automédication occasionnelle par antalgique de pallier 1 (paracétamol / AINS), les céphalées de tensions fréquentes et chroniques sont responsables pour le patient d'un risque d'abus médicamenteux et pour la société d'un coût important. Le coût par patient est estimé en Europe à environ 3 864 € pour les abus médicamenteux et à 303 € pour les céphalées de tension, dus à 92% à une baisse de productivité et aux arrêts de travail. (2)

Afin de soulager le patient et prévenir le risque de surconsommation d'antalgiques, un traitement de fond est préconisé pour ceux souffrant de céphalées de tension fréquentes ou chroniques.

Les mécanismes de la céphalée de tension restent mal connus. L'étiologie la plus probable est une contraction anormale de certains muscles (trigger point) induisant une hypoxie musculaire et une libération des médiateurs de la douleur. L'exposition répétée et prolongée à ces neuromédiateurs modifierait la sensibilité à la douleur du Système Nerveux Périphérique (SNP) puis du Système Nerveux Central (SNC) qui induirait les céphalées de tension (3). Compte-tenu de son action antalgique sur le SNC, l'amitriptyline est la molécule préconisée en première intention pour le traitement des céphalées de tension en accompagnement d'une prise en charge paramédicale, par la société suisse pour l'étude des céphalées (2019) (4), European Federation of Neurological Societies (EFNS) (2010) (5), et l'Agence Nationale d'Accréditation et d'Evaluation en Santé (ANAES) (2004) (6).

L'amitriptyline découverte en 1960 est un antidépresseur tricyclique avec une action antalgique distincte. Cette molécule commercialisée pour la première fois aux Etats Unis en 1961 (7) présente plusieurs propriétés pharmacologiques distinctes impliquant divers mécanismes d'actions :

Inhibition de la recapture présynaptique de la sérotonine et de la noradrénaline assurant sa fonction antidépressive (AMM renouvelée en 2021, Service Medical Rendu jugé important en 2016)

Action histaminique assurant une sédation (pas d'AMM)

Au niveau cérébral et médullaire, inhibiteur des canaux NMDA et sodiques et recapture de la noradrénaline assurant son action analgésique (AMM douleur neuropathique avec SMR jugé important 2021, AMM pour migraine et céphalées de tension 2021)

Action anticholinergique ayant probablement été le mécanisme d'action pour le traitement de l'énurésie nocturne isolée chez l'enfant (SMR jugé insuffisant à cause du faible bénéfice par rapports à ses effets indésirables depuis 2016 et son retrait de l'AMM depuis) (8-9). Cet effet est responsable de la majorité de ses effets indésirables fréquents (touchant plus de 1 patient sur 10) au premier rang desquels : la sécheresse buccale.

Cependant il n'y a aucune méta-analyse dédiée uniquement à l'amitriptyline dans la céphalée de tension contre placebo alors que la molécule est souvent utilisée comme traitement de référence pour d'autres molécules. Il existe une revue de la littérature sur le sujet datant de 2008 par Torrente Castells et al intitulée "Use of amitriptyline for the treatment of chronic tension-type headache". Cependant nous constatons qu'elle utilise une méthodologie différente et qu'elle inclut entre autres des études ne comportant pas de placebo en comparaison (10). L'amitriptyline étant un traitement de référence dans la prise en charge des patients atteints de céphalées de tension et dans les essais cliniques, notre travail visera à vérifier d'une part la qualité scientifique des études portant sur son efficacité, et d'autre part à objectiver l'amélioration des céphalées.

2. Objectifs :

Notre travail vise à vérifier le niveau de preuve de l'amitriptyline pour le traitement des patients adultes atteints de céphalées de tension, par une revue systématique de la littérature avec une méta-analyse des essais randomisés contre placebo, à travers 3 critères de jugements : la durée, la fréquence et l'intensité des céphalées.

II- MATERIELS ET METHODES :

1. Protocole et enregistrement :

Le protocole de recherche de cette revue systématique de la littérature a été enregistré le 21 janvier 2022 sur la plateforme OSF sous le numéro d'immatriculation : 10.17605/OSF.IO/M9QS2

2. Critères d'éligibilité :

Les critères d'inclusion retenus pour cette revue de la littérature sont les suivants :

Les études doivent être des essais prospectifs, randomisés, en double aveugle. L'amitriptyline doit être testée contre un placebo. Tous les schémas d'intervention sont éligibles aussi bien les études en parallèle qu'en cross-over. La population étudiée doit être une population adulte de plus de 18 ans.

Le critère de jugement principal de notre étude est la douleur, évaluée de plusieurs manière (intensité, durée et fréquence). Toutes les études évaluant ces paramètres en critère principal ou secondaire sont éligibles. Il n'y a pas de restriction quant à la posologie du traitement, sa durée ou l'année de publication de l'étude. Les articles doivent par ailleurs être rédigés en anglais ou en français.

3. Sources d'information :

La recherche des articles éligibles a été réalisée sur trois bases de données bibliographiques scientifiques, toutes consultées en ligne :

- ClinicalTrials.gov
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Medline (PubMed)

Dans un second temps, les études retenues par les auteurs ayant précédemment réalisé des revues systématiques de la littérature et/ou méta-analyses dans le même domaine de recherche

ont été consultées, afin d'assurer une collecte optimale des articles éligibles. Une fois la recherche bibliographique terminée, une vérification des bibliographies des articles a été réalisée afin de n'en omettre aucun.

La recherche bibliographique s'est déroulée jusqu'à janvier 2020. Une veille bibliographique a été réalisée jusqu'en septembre 2021.

4. Recherche :

Le mot clé MeSH employé pour la recherche bibliographique est « AMITRIPTYLINE ». Le filtre « randomized controlled trial » a été ajouté pour la recherche d'articles dans les bases de données Cochrane CENTRAL et Medline (PubMed).

5. Sélection des études :

Le logiciel ZOTERO ® a été utilisé pour la sauvegarde et le tri des références bibliographiques. Les auteurs : Valérie GUICHARDANT (V.G.) et Houdna LAIB (H.L.) ont réalisé indépendamment une recherche bibliographique exhaustive dans les 3 bases de données ci-dessus.

Dans un premier temps les auteurs ont utilisé "amitriptyline" comme mot clé de recherche dans PubMed et Cochrane avec le filtre "essai clinique randomisé". Ils ont sélectionné par la lecture des titres toutes les études concernant les céphalées de tension chroniques. Pour chaque article dont le titre a semblé pertinent, le résumé a été lu. Ceci a permis de déterminer si l'essai clinique remplit les critères d'inclusion. Si le résumé d'un article ne suffisait pas pour statuer, ce dernier a été lu en intégralité. L'éligibilité des articles présents dans les méta-analyses et les revues systématiques publiées dans Pubmed et Cochrane, ont été analysés de la même façon.

A la suite de ce travail, les deux auteurs ont confronté leur bibliographie. Les articles qu'elles ont toutes deux retenus ont été inclus ; les autres ont fait l'objet d'une réévaluation commune. En cas de désaccord persistant, un troisième intervenant indépendant a été sollicité pour trancher, le docteur Yves BARO (Y.B.) ou le professeur Rémy BOUSSAGEON (R.B.).

Une fois ces articles sélectionnés, les doublons d'une même étude ont été identifiés et éliminés. Les textes intégraux des études sélectionnées ont été récupérés. Les auteurs ont alors analysé les bibliographies des articles sélectionnés, afin d'être exhaustives dans leur recherche bibliographique.

Le processus de sélection des articles est rapporté à l'aide d'un diagramme de flux selon les lignes directrices du protocole PRISMA 2020 (11).

6. Extraction des données :

Les données des études ont été extraites indépendamment par les deux auteurs (V.G. et H.L.) selon les caractéristiques suivantes :

Méthode de l'étude : Type d'étude : parallèle ou cross-over ; durée de l'étude : durée totale pour les essais en parallèles, période de traitements et de wash-out pour les essais en cross over ; type d'analyse réalisée (ITT, ITTm ou per protocole).

Population étudiée : nombre de participants randomisés ; nombre de participants exclus et perdus de vue ; nombre de participants en fin d'intervention ; critères d'inclusion et d'exclusion ;

Intervention de l'étude : nom des molécules testées dont amitriptyline et placebo ; posologies administrées pendant l'essai ; thérapies non médicamenteuses testées.

Critères de jugement : critères principaux et secondaires

Résultats de l'étude : significatifs ou non pour l'amitriptyline versus placebo

Notes : utilisation ou non d'une méthode d'ajustement de la significativité en cas de critères de jugement multiples, conflits d'intérêts déclarés

7. Risque de biais inhérent à chacune des études

L'outil « Revised Cochrane risk-of-bias tool for randomized trials (RoB 2) » (12) a été appliqué indépendamment par les deux auteurs (V.G. et H.L.) pour chaque étude, afin d'évaluer sa qualité méthodologique.

Ils ont ensuite confronté leurs analyses de biais, et en cas de différence, le biais a été réanalysé. En cas de non-aboutissement à un consensus, les auteurs ont soumis l'analyse à un troisième intervenant indépendant pour statuer : le docteur Yves BARO (Y.B.) ou le professeur Rémy BOUSSAGEON (R.B.).

L'outil RoB 2 propose d'analyser la robustesse scientifique de chaque étude au travers de 5 catégories de biais, évalués chacune indépendamment en trois niveaux : risque faible, quelques inquiétudes et risque important :

- Biais résultants du processus de randomisation
- Biais résultants des écarts aux interventions prévues
- Biais résultants des données manquantes
- Biais résultants de la mesure du critère de jugement
- Biais résultants de la sélection des résultats rapportés

L'étude est considérée comme étant à faible risque de biais si chacune des 5 catégories est jugée à « faible risque de biais ». Elle est considérée comme présentant quelques inquiétudes si 4 catégories sur 5 sont à « faible risque de biais », mais que 1 catégorie sur 5 pose « quelques inquiétudes ». Elle est enfin considérée comme étant à haut risque de biais si au moins une catégorie de biais est jugée à « haut risque » ou si plusieurs catégories présentent « quelques inquiétudes ».

L'application de l'outil RoB 2 sur les études incluses a suscité de nombreuses interrogations et parfois l'outil n'a pas apporté de réponse adéquate à ces questions. Les deux auteurs ont alors instauré avec l'accord du Pr. Boussageon des règles générales permettant de statuer plus facilement.

Premièrement, pour les études présentant des critères de jugement multiples, des méthodes d'ajustement permettant le contrôle du risque d'erreur de type I seront recherchées. En l'absence d'ajustement du risque alpha, l'analyse sera jugée inadéquate même si une différence statistique est retrouvée.

Deuxièmement, pour les différentes catégories de biais :

Concernant le “Biais résultats du processus de randomisation” :

- Il a été statué que lorsqu’une étude est de grande envergure, par exemple réalisée dans plusieurs centres universitaires ou que l’équipe d’investigation a réalisé plusieurs publications précisant dans l’une d’elle une méthode de randomisation valide ; bien qu’il n’existe pas de détails concernant la séquence d’allocation ni de sa dissimulation dans l’étude analysée ; la réponse à l’item sera “probablement oui” plutôt que “absence d’informations”.
- Lorsqu’il s’agit d’études en crossover, il a été décidé que lorsque les auteurs de l’essai ne présentent pas les caractéristiques de base de la population initiale la réponse à l’item sera “absence d’information” plutôt que “probablement non”.

Concernant le “Biais résultants des écarts aux interventions prévues” :

- Lorsque l’étude exclut uniquement les perdus de vue, celle-ci sera considérée comme utilisant une analyse en intention de traiter modifiée (ITTm). Mais si une étude exclut des participants pour tout autre raison, alors celle-ci sera considérée comme utilisant une analyse en per protocole.
- Le double aveugle est considéré comme levé lorsqu’il existe plus de 50% d’effets indésirables spécifiques à l’amitriptyline définis par somnolence ou sécheresse buccale dans le groupe traité comparé au groupe placebo. (8-9)
- Lorsque dans les essais cliniques les effets indésirables spécifiques à l’amitriptyline ne sont pas rapportés, les auteurs (V.G. et H.L.) ont considéré qu’il était difficile de statuer sur le maintien du double aveugle et dans ce cas la réponse à l’item sera “absence d’information”.
- Une analyse en ITTm est jugée comme inadéquate s’il existe plus de 5% de données manquantes.
- Lorsque l’analyse employée est jugée inadéquate, il a été décidé d’estimer systématiquement qu’il existe un risque d’une répercussion sur le résultat.

Concernant le “ Biais résultants des données manquantes ” :

- Le seuil des données manquantes est fixé à 5%.
- Lorsque les causes de données manquantes sont mal précisées, l'estimation de leur possible impact sur la vraie valeur sera jugée “ absence d'information”.

Concernant le “ Biais résultants de la mesure du critère de jugement ” :

- Lorsque les critères de jugement sont considérés comme étant subjectifs, et que le double aveugle est considéré comme levé alors l'estimation d'une possible influence de l'évaluation par la connaissance du groupe d'intervention est jugée “probablement oui”.
- Il a été décidé de ne pas se positionner concernant l'influence de la connaissance de l'intervention reçue sur l'évaluation des résultats avec un jugement porté “absence d'informations”.

Concernant le “Biais dans la sélection des résultats rapportés” :

- Lorsqu'il n'est pas fait mention que le protocole de l'étude a été enregistré préalablement, il a été statué “absence d'informations”.
- Lorsque les auteurs de l'essai annoncent avoir trouvé une différence significative sans donner le résultat il a été décidé qu'il y avait un risque de sélection de résultat avec un jugement porté : “probablement oui”.

Concernant le risque d'inflation du risque alpha dans le cas de multiplicité des tests :

- Lorsque les auteurs rapportent avoir réalisé un ajustement lors d'une précédente étude, il a été statué que pour l'étude en question il est considéré que l'ajustement a également été réalisé.

Etant donné que nous avons statué :

- La levée du double aveugle en cas d'absence d'information sur les effets indésirables spécifiques de l'amitriptyline
- Un seuil de données manquante et d'acceptabilité de perdus de vu analyse en ITTm à 5%

- Une réponse “ absence d’information “ si les auteurs ne déclarent pas avoir enregistré leur protocole au préalable

Lorsque ces décisions impactent fortement l'évaluation du risque de biais de chaque catégorie, il a été décidé avec Pr. Boussageon de réaliser une étude de sensibilité méthodologique lorsque cela est pertinent. Ainsi, pour chaque article scientifique analysé, nous regarderons si son niveau global de preuve est impacté par UNE SEULE de ses trois décisions.

Si c'est le cas, l'article sera réanalysé avec :

- Pour la non mention des effets indésirables une estimation d'absence d'impact sur le double aveugle
- Pour les données manquantes et le seuil d'acceptabilité de perdu de vue d'ITTm à 10 et 15%
- Une réponse “ probablement oui “ si les auteurs ne fournissent aucune information sur l'enregistrement du protocole

8. Quantification des résultats :

Les auteurs ont retenu la douleur comme critère de jugement principal mesurée de plusieurs manières : fréquence, intensité et durée. La mesure de ces résultats est dans la majorité des études réalisée à l'aide d'une échelle visuelle analogique. Les résultats sont de ce fait présentés sous forme de moyennes $\pm$ l'écart-type (SD).

9. Synthèse des résultats :

L'ensemble des études incluses sera analysé, afin de constituer une synthèse qualitative.

Si suffisamment de résultats sont éligibles, sous forme de moyenne avec écart type, la synthèse quantitative sera également réalisée, à l'aide du logiciel Review Manager version 5.4 (RevMan®) selon les préconisations de la méthodologie REB (13).

L'analyse quantitative sera conditionnée par la gestion des données manquantes. Pour ce faire, les auteurs des ECR et des méta-analyses ont été sollicités par mail, afin qu'ils nous transmettent leurs recueils de données. Des relances hebdomadaires seront réalisées sur la période d'un mois. Les articles pour lesquels la moyenne $\pm$ l'écart type d'aucun des trois critères de jugement ne pourront être obtenu seront exclus de l'analyse quantitative.

Selon la diversité des échelles de mesures de nos critères de jugement dans les articles, des différences de moyennes (si besoin standardisées) seront calculées à partir des résultats ainsi que leurs intervalles de confiance à 95%. Le risque alpha sera fixé à 0,05. L'hétérogénéité des résultats des essais sera par la suite évaluée grâce à un test statistique de Chi2. Un modèle d'effet fixe sera utilisé préférentiellement. En cas d'hétérogénéité forte ($I^2 > 50\%$) un modèle d'effet aléatoire sera utilisé.

Dans la mesure du possible, deux méta-analyses seront réalisées pour pouvoir répondre au protocole REB (13) et parvenir à une conclusion :

- Une méta-analyse principale, qui ne prendra en considération que les résultats d'essais cliniques randomisés (ECR) de haute qualité. Le résultat de cette méta-analyse sera considéré comme positif si les résultats sont statistiquement significatifs ($p < 0,05$). Le nombre d'ECR concluants selon l'analyse du RoB 2, l'hétérogénéité et le biais de publication entre les ECR seront évalués.
- Une autre méta-analyse dite "de sensibilité" sera réalisée incluant tous les ECR pour lesquels les résultats sont analysables, quelle que soit leur qualité, permettant de vérifier la robustesse des résultats de la méta-analyse de base.

10. Risque de biais transversal aux études et évaluation du niveau de preuve :

Le niveau de preuve des résultats de la méta-analyse sera évalué selon l'approche GRADE (Grading of recommendations assessment, development and evaluation)(14).

Le niveau de preuve des résultats issus d'une méta-analyse d'ECR est en principe le plus élevé, mais il peut être diminué dans plusieurs situations : en cas de risque de biais dans les ECR, d'imprécision (intervalle de confiance trop large), de forte hétérogénéité (résultats contradictoires), de preuve "indirecte" (population, intervention, critère ou comparateur non totalement adapté) ou de présence d'un biais de publication. Ainsi selon les caractéristiques remplies le niveau de preuve pourra être défini comme "élevé", "modéré" ou "faible".

Dans le cadre de la méthodologie REB une conclusion finale sera proposée en fonction de la configuration de l'étude et de plusieurs critères :

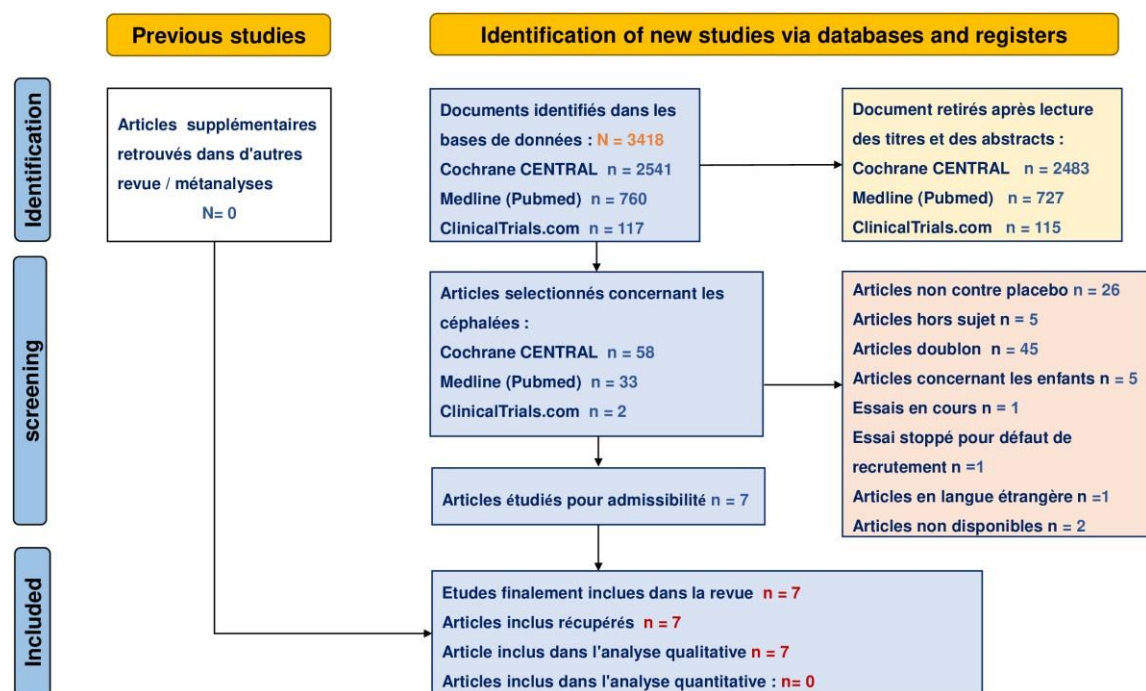
- *Si une méta-analyse principale trouve des résultats positifs* : trois conclusions sont possibles selon 4 critères (Nombre d'ECR concluants, Niveau d'hétérogénéité, Biais de publication et Résultats de la méta-analyse de sensibilité).

- Si une méta-analyse principale n'est pas possible mais qu'une analyse de sensibilité est possible : trois conclusions sont possibles en fonction de 3 critères (nombre d'ECR concluants, Biais de publication, Résultats de la méta-analyse de sensibilité).
- Si ni l'analyse principale ni l'analyse de sensibilité ne sont possibles : trois conclusions sont possibles en fonction de 2 critères (nombre d'ECR concluants et biais de publication).

Les trois conclusions possibles sont :

- Preuve solide
- Preuve à confirmer
- Absence de preuve

III- SELECTION DES ETUDES :



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

La recherche avec le seul mot clé « amitriptyline » et l'usage dans PubMed du filtre « essai clinique randomisé » a mis en évidence 3418 articles dont 2541 dans Cochrane, 760 dans Medline et 117 dans Clinical Trials.

Les articles dont les titres et abstracts ne remplissaient pas les critères d'inclusion récapitulés ci-dessous ont été exclus :

- Essai clinique randomisé
- En double aveugle
- Comparant l'amitriptyline à un placebo
- Rédiger en français ou en anglais
- Dont le texte intégral est disponible
- Ayant comme objectif principal ou secondaire l'évaluation de la douleur selon différentes modalités : la fréquence, la durée ou l'intensité.

93 articles correspondaient par leur titre et abstracts aux critères d'inclusion : 58 articles dans Cochrane, 33 dans Medline et 2 dans Clinical Trials.

Parmi ces articles, il a été impossible de récupérer le texte intégral pour 2 d'entre eux, 1 était rédigé en langue étrangère, 5 se sont révélés être hors sujet, 25 n'évaluaient pas l'amitriptyline contre un placebo, 5 articles concernaient une population pédiatrique, 1 essai est en cours et 1 autre a été stoppé pour défaut de recrutement. Par ailleurs 45 étaient des doublons.

Nous avons pu finalement inclure 7 articles pour notre analyse :

[A - J.W Lance 1964](#) : Treatment of chronic tension headache (15).

[B - Gobel 1994](#) : Chronic tension-type headache: amitriptyline reduces clinical headache-duration and experimental pain sensitivity but does not alter pericranial muscle activity readings (16).

[C - Pfaffenrath 1994](#) : Efficacy and tolerability of amitriptylinexide in the treatment of chronic tension-type headache: a multi-centre controlled study (17).

[D - Brentsen 1996](#) : A non-selective (amitriptyline), but not a selective citalopram), serotonin reuptake inhibitor is effective in the prophylactic treatment of chronic tension-type headache (18).

[E - Diamond 1998](#) : Chronic tension headache--treated with amitriptyline a double-blind study (19).

[F - Brendtsen 2000](#) : Amitriptyline reduces myofascial tenderness in patients with chronic tension-type headache (20).

[G - Ashina et Brendtsen 2004](#) : Analgesic effect of amitriptyline in chronic tension-type headache is not directly related to serotonin reuptake inhibition (21).

Les bibliographies des méta-analyses sur le traitement des céphalées de tension et des articles sélectionnés n'ont pas permis de mettre en évidence d'autres essais cliniques éligibles à notre étude.

IV- RÉSULTATS

1. Caractéristiques des études sélectionnées :

Les tableaux disponibles en annexe détaillent les caractéristiques de chaque étude incluse.

Nous avons inclus sept études dans notre analyse qualitative : Lance 1964, Gobel 1994, Pfaffenrath 1994, Brendtsen 1996, Diamond 1998, Brendtsen 2000, Ashina et Brendtsen 2004. Ces études ont été publiées entre 1964 et 2004.

Toutes les études sont monocentriques hormis l'étude de Pfaffenrath 1994 réalisée dans 3 pays sur un total de sept centres.

Parmi ces sept études, quatre suivent un schéma d'étude en crossover (J.W Lance 1964, Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004) et trois suivent un schéma d'étude en parallèle (Gobel 1994, Pfaffenrath 1994 et Diamond 1998).

Toutes les études testent l'amitriptyline contre un placebo inactif.

Quatre études sont composées de trois bras d'essai (Brendtsen 1996, Diamond 1998, Brendtsen 2000, Ashina et Brendtsen 2004,). Trois études testent l'amitriptyline contre la molécule citalopram (Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004) et une étude teste l'amitriptyline à deux dosages différents 10 mg et 25 mg versus placebo (Diamond 1998).

Deux études sont à deux bras d'essai testant l'amitriptyline contre un placebo (Göbel 1994, Pfaffenrath 1994). L'étude de J.W Lance 1964 présente deux essais : le premier à deux bras teste l'amitriptyline et un placebo et un second essai composé de douze bras (des sédatifs (amylobarbitone), des vasodilatateurs, la combinaison d'un sédatif et d'un médicament vasodilatateur, du "ballergal " un mélange de phénobarbital, d'ergotamine et de belladone, du

méthysergide un dérivé d'acide lysergique, du citrate d'orphénadrine, du chlórdiazépoxide, du diazépam, de l'amitriptyline, de l'Imipramine, du "parstelin" un mélange de tranlycypromine et trifluopérazine et un placebo).

La durée de l'étude pour les groupes parallèles est de 4 semaines pour Diamond 1998, de 6 semaines pour Göbel 1994 et 24 semaines pour Pfaffenrath 1994. Pour les études en crossover, les études Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004 ont mené leurs études sur une durée totale de 32 semaines avec une période initiale de 4 semaines, trois périodes de traitements de 8 semaines séparés par deux périodes de wash out d'une durée de deux semaines. L'étude de Lance 1964, également en cross-over a mené deux essais : le premier d'une durée totale de 12 mois sans période initiale ni période de wash out, le second d'une période totale de 2 mois sans période initiale ni période de wash out.

Le nombre de participants randomisés varie entre 27 et 280 patients selon les études, pour un total de 792 patients. L'âge des patients est compris entre 18 et 75 ans.

Pour être inclus les patients devaient présenter des céphalées de tension depuis plus d'un an pour une étude (Lance 1964) et au moins 15 jours par mois durant les 6 derniers mois pour cinq études (Göbel 1994, Pfaffenrath 1994, Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004). L'étude de Diamond 1998 ne définit pas de durée minimale de céphalées. L'étude de Diamond 1998 inclut par ailleurs uniquement des patients présentant de l'anxiété ou une dépression associée aux céphalées.

Les critères d'exclusion varient selon les études avec pour Lance 1964 l'exclusion uniquement des patients présentant des céphalées peu fréquentes ou occasionnelles, tandis que pour Diamond 1998 les patients ne présentant pas de céphalées définies comme légères/modérées/sévères et les patients ne présentant aucun des symptômes ciblés étaient exclus de l'étude.

Pour les six autres études plusieurs critères d'exclusion, souvent similaires, sont cités, parmi ceux-ci :

- Une grossesse, un projet de grossesse ou un allaitement (Göbel 1994, Pfaffenrath 1994). Les femmes en âge de procréer ont dû utiliser des mesures contraceptives adéquates tout au long de l'étude pour Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004.
- La participation à une étude au cours des 3 derniers mois (Pfaffenrath 1994), toute participation antérieure à une étude (Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004)

- Les contre-indications à l'utilisation de l'amitriptyline (épilepsie, glaucome, HTA, dysfonction prostatique, troubles cardiaques ou hépatiques) (Göbel 1994, Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004)
- Une hypersensibilité connue à l'amitriptyline ou une allergie connue (toutes les études)
- Une addiction médicamenteuse/toxicomanie/ abus médicamenteux (Göbel 1994, Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004). La prise d'un antalgique simple ou mixte, de tartrate d'ergotamine ou de tartrate de di-hydroergotamine, d'acide acétylsalicylique et/ou de paracétamol ou de codéine plus de 10 jours par mois, la prise d'autres antidépresseurs, neuroleptiques, tranquillisants régulièrement (Pfaffenrath 1994). La prise de médicaments psychoactifs dans les 6 mois précédant l'essai et pendant l'essai pour Göbel 1994 ; dans les 3 mois précédant la phase initiale pour Pfaffenrath 1994 et dans les 4 semaines pour les inhibiteurs de la MAO. Le mésusage d'analgésiques simples (correspondant à plus de 2 g d'aspirine par jour), une prise régulière d'opiacés ou de benzodiazépines pour Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004.
- Un antécédent de traitement par un antidépresseur (Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004)
- Des antécédents de maladies psychiatriques (Göbel 1994), les affections psychiatriques graves y compris la dépression avec un score de dépression de Hamilton supérieur ou égal à 17 (Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004)
- Les autres formes de céphalées primaires ou symptomatiques dont les épisodes de migraines (Göbel 1994, Pfaffenrath 1994, Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004)
- La non-participation aux rendez-vous de suivi (Göbel 1994) une mauvaise observance suspectée (Pfaffenrath 1994)

Six études sur les sept ont un critère de jugement principal évaluant les céphalées selon différentes modalités (Lance 1964, Göbel 1994, Pfaffenrath 1994, Brendtsen 1996, Diamond 1998, Brendtsen 2000, Ashina et Brendtsen 2004).

L'étude de Diamond 1998 ne précise pas si l'évaluation des céphalées est un critère de jugement principal ou secondaire.

Deux études évaluent les céphalées selon la durée des céphalées en critère principal (Göbel 1994 et Pfaffenrath 1994), l'étude de Brendtsen 1996 évalue également la durée des céphalées mais en critère secondaire.

L'étude de Pfaffenrath 1994 évalue les céphalées selon l'intensité des céphalées en critère principal, les études de Brendtsen 1996 et de Brendtsen 2000 évaluent également l'intensité des céphalées mais en critère secondaire.

L'étude de Pfaffenrath 1994 évalue les céphalées selon la fréquence des céphalées en critère principal, l'étude de Brendtsen 1996 évalue également la fréquence des céphalées mais en critère secondaire.

L'étude de Lance 1964 évalue les céphalées selon un critère de jugement principal composite : “ l'amélioration des symptômes selon la fréquence ou l'intensité”.

L'étude de Diamond 1998 évalue les céphalées de manière indifférenciée sans précision concernant le caractère principal ou secondaire de ce critère de jugement.

Trois études évaluent les céphalées selon une courbe AUC calculée à partir du produit durée multiplié par l'intensité des céphalées en critère de jugement principal pour Brendtsen 1996 et en critère de jugement secondaire pour Brendtsen 2000 et Ashina et Brendtsen 2004.

L'étude de Pfaffenrath 1994 évalue également en critère de jugement principal les céphalées selon le pourcentage de patients améliorés par le calcul de :

- La fréquence multipliée par la durée des céphalées
- La durée multipliée par la fréquence x l'intensité
- La réduction d'au moins 50 % du produit durée multiplié par la fréquence Et au moins une réduction de 50 % dans l'intensité des céphalées

D'autres critères de jugement principaux ont été évalués dans les études :

- L'enregistrement des effets indésirables avec le questionnaire FSUCL (Pfaffenrath 1994)
- La sensibilité myofasciale (Brendtsen 2000)
- Le seuil de douleur à la pression (Brendtsen 2000)
- Le seuil de douleur électrique (Brendtsen 2000)
- Le taux de plaquettes 5 HT (Ashina et Brendtsen 2004)

- L'activité EMG, la suppression extéroceptive de l'activité du muscle temporal, la variation négative du contingent (CNV) et l'expérience de la sensibilité à la douleur (Göbel 1994)

Les autres critères de jugements secondaires évalués dans les études sont :

- L'auto-évaluation des patients par le "Self Report Symptom Inventory 90 items-R (Pfaffenrath 1994)
- Le taux 5-HT plaquettaire dans le plasma (Ashina et Brendtsen 2004)
- Le nombre de doses d'analgésiques, équivalent à 500 mg d'aspirine, pris au cours des quatre dernières semaines de chaque période de traitement ((Brendtsen 2000 ,Ashina et Brendtsen 2004)
- Le nombre d'effets indésirables au cours de chaque période de traitement ((Brendtsen 2000, Ashina et Brendtsen 2004)

L'étude de Diamond 1998 explore plusieurs critères de jugements sans précision concernant le caractère principal ou secondaires de ceux-ci :

- L'anxiété, l'humeur dépressive, les plaintes psychosomatiques, l'insomnie, l'apathie, la fatigue, la baisse de l'efficacité au travail
- La réponse globale hebdomadaire au traitement par le patient et l'évaluation globale faite par le médecin
- Le poids des patients, la tension artérielle, le pouls, la numération sanguine, une analyse d'urines, de l'azote, d'urée dans le sang (BUN) et des transaminases

Trois études sont considérées comme “confirmatoire” pour leur critère de jugement principal car elles ont des critères de jugement principaux uniques (Lance 1964, Göbel 1994 et Brendtsen 1996). Trois études sur sept ont des critères de jugement principaux multiples mais toutes utilisent une méthode d'ajustement du risque de première espèce et leurs résultats sont donc également considérés comme “ confirmatoires” pour leurs critères de jugements principaux. (Pfaffenrath 1994, Brendtsen 2000, Ashina et Brendtsen 2004). L'étude de Diamond 1998 est considérée quant à elle comme étant “exploratoire” devant la multiplicité des critères de jugements, sans précision apportée sur le caractère principal ou secondaire, ni aucune information concernant l'utilisation d'une méthode d'ajustement du risque alpha.

Trois études ont réalisé leur analyse de résultat selon une “intention de traiter” (Lance 1964, Pfaffenrath 1994 et Diamond 1998), trois autres selon une “intention de traiter modifiée” (Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004) et une étude selon “une per protocole” (Göbel 1994).

Le pourcentage de perdus de vue varie entre 5.12 et 32% selon les études. L’étude de Lance 1964 ne fournit aucune information concernant la proportion de patients perdus de vue.

Trois études indiquent avoir été soutenues par la fondation Lundbeck (Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004) et l’une par la fondation pharmaceutique 1991 (Ashina et Brendtsen 2004).

2. Risque de biais relatif aux études

Les tableaux disponibles en annexe détaillent le risque de biais pour chaque étude selon l’outil RoB2 et selon l’inflation du risque alpha.

Les sept études incluses ont un risque de biais global élevé.

Nom l'auteur principal et année de publication	Biais résultant du processus de randomisation	Biais dû à l'effet de période et à l'effet carry over	Biais résultant du aux écarts aux interventions prévues	Biais résultant des données manquantes	Biais résultant de la mesure du critère de jugement	Biais résultant de la sélection des résultats rapportés	Biais résultant de l'inflation du risque alpha	Evaluation globale
Lance et Curran 1964								Risque élevé
Gobel 1994								Risque élevé
Pfaffenrath 1994								Risque élevé
Brendtsen 1996								Risque élevé
Diamond 1998								Risque élevé
Brendtsen 2000								Risque élevé
Ashina et Brendtsen 2004								Risque élevé

Vert : risque faible ; Orange : risque modéré ; Rouge : risque élevé

Le graphique ci-dessous présente pour chaque domaine de biais, le pourcentage d'études jugées à risque « faible » ou présentant « quelques inquiétudes » ou à risque « élevé » (Figure 1).

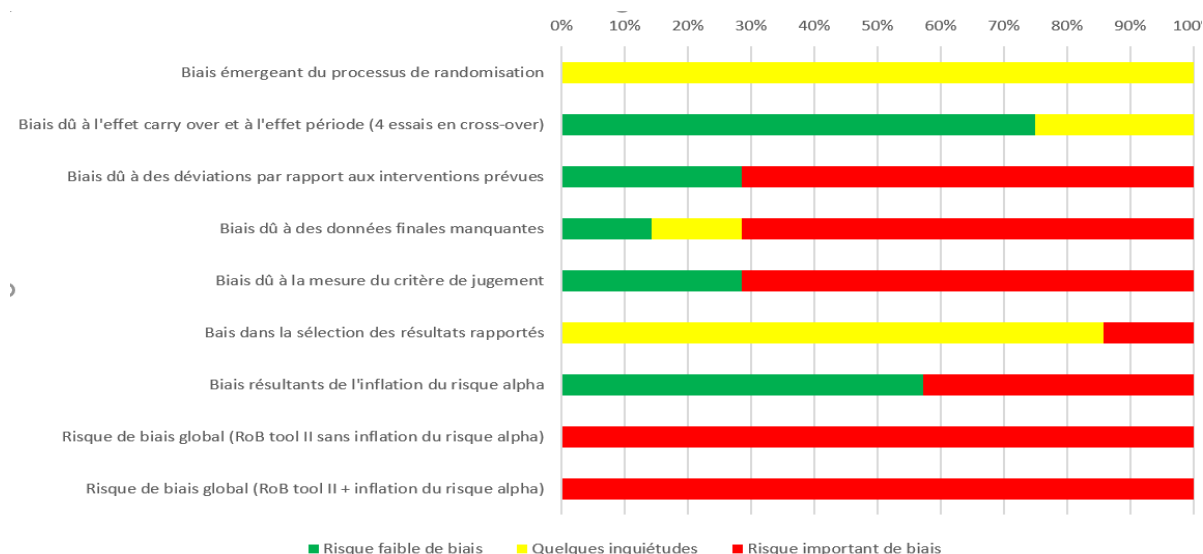


Figure 1 : Evaluation par domaine de biais du pourcentage d'études présentant un risque du risque de biais estimé « faible » ou présentant « quelques inquiétudes » ou « élevé »

Le graphique (Figure 1) montre pour chaque domaine de biais le pourcentage d'études présentant un risque jugé élevé. Les 3 domaines de biais estimés à risque élevé les plus fréquents sont : les biais liés aux déviations par rapport aux interventions prévues, les biais liés aux données manquantes et les biais liés aux mesures du critère de jugement (71.43% des études). Ensuite viennent les biais liés à la sélection des résultats rapportés (14.29% des études). L'outil RoB tool II de Cochrane ne tient pas compte du biais résultant de l'inflation du risque alpha contrairement à la méthodologie REB. Il est estimé élevé pour 42.86% des études.

a) Les biais dus au processus de randomisation :

Pour les 7 études, le risque de biais dans ce domaine est jugé comme étant modéré.

Seule l'étude de Pfaffenrath 1994 détaille le processus de randomisation avec une séquence d'allocation jugée comme probablement aléatoire. Toutes les autres études ne détaillent pas le processus de randomisation, Il n'y a ainsi pas d'informations fournies sur la séquence d'allocation. Concernant la dissimulation de la séquence d'allocation, seule l'étude de Lance 1964 apporte des éléments permettant de supposer que celle-ci ai été réalisée. Pour toutes les autres études il n'y a pas d'informations fournies sur une dissimulation de séquence d'allocation. Les caractéristiques de base des patients entre les groupes d'intervention au début de l'étude ont été comparées sans mettre en évidence de différence qui ferait évoquer un problème au niveau du processus de randomisation pour l'étude de Pfaffenrath 1994, De

Brendtsen 1996 et Brendtsen 2000. Pour les études de Göbel 1994 et Diamond 1998, les caractéristiques de la population finale sont présentées et non de la population initiale. Pour les études de Lance 1964 et Ashina et Brendtsen 2004 il n'y a pas d'information donnée sur les caractéristiques de la population étudiée.

b) Les biais dus à l'effet de période et à l'effet carry-over :

4 études sont des essais en cross over : Lance 1964, Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004, L'évaluation du risque de biais dans ce domaine a été jugé faible pour Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004 tandis qu'elle a été jugée à risque modéré pour Lance 1964. Dans les 3 études à faible risque le rapport des participants dans chaque bras de chaque période de traitement est de 1 pour 1. Dans l'étude de Lance 1964 le nombre des participants dans chaque bras semble être aléatoire au cours de l'étude. Il varie en fonction de la réponse au traitement attribué. En effet si le patient répond au traitement il poursuit l'étude avec celui-ci mais s'il n'y a pas de réponse le patient se voit alors attribué un nouveau traitement et change ainsi de bras. Dans les 3 études à faible risque il n'y a pas eu d'effet "carry-over" identifié. Pour l'étude de Lance il n'y a pas de période de "wash out" mais en tenant compte du fait que l'amitriptyline a une demi-vie d'élimination plasmatique comprise entre 22 et 40 heures selon les individus et que les mesures sont réalisées 4 semaines après un changement de molécule, nous considérons qu'il n'y a pas d'effet "carry-over".

c) Les biais dus aux écarts aux interventions prévues :

Pour ce domaine, nous avons jugé que 5 études présentent un risque de biais élevé. Il s'agit des études de Göbel 1994, Brendtsen 1996, Diamond 1998, Brendtsen 2000 et Ashina et Brendtsen 2004.

Pour les études de Göbel 1994, Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004, nous avons considéré que le double aveugle était levé compte tenu de l'occurrence au moins deux fois plus fréquente dans les groupes testant l'amitriptyline, d'effets indésirables spécifiques de la molécule tels qu'une sécheresse buccale et/ou une somnolence (8-9). Pour l'étude de Diamond 1998 nous avons considéré que le double aveugle était probablement levé devant un nombre de patients présentant des effets secondaires huit fois plus important pour les patients traités avec de l'amitriptyline sans préciser plus de détails sur la nature de ces effets secondaires.

L'étude de Göbel 1994 est la seule étude utilisant une analyse en "per protocole". Pour cette étude, le manque d'adhésion à l'intervention a été jugé comme étant important avec notamment 21 patients qui n'ont pas respecté le protocole de l'étude. Il n'est pas fait mention par ailleurs de l'utilisation d'une analyse appropriée pour estimer l'effet de l'adhésion.

Les études de Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004 analysent leurs résultats selon une "intention de traiter modifiée". Cette analyse est cependant jugée comme inappropriée pour ces 4 études car le pourcentage de patients perdus de vue dépassait les 5%. Pour l'étude de Diamond 1998, plusieurs analyses ont été réalisées selon les résultats mesurés soit une "ITT" soit une "ITT modifiée" ont été appliquées. L'analyse en "ITTm" réalisée est considérée comme inappropriée compte tenu du taux important de données manquantes (32%). Pour toutes les études il n'y a pas eu a priori d'écarts par rapport aux protocoles respectifs prévus.

2 études sont à faible risque de biais pour ce domaine : Lance 1964 et Pfaffenrath 1994.

Pour ces 2 essais cliniques, nous avons considéré que le double aveugle a été maintenu. Il n'a pas été constaté d'écarts aux interventions prévues par leurs protocoles respectifs. Les deux études utilisent une analyse "en intention de traiter" qui est jugée comme étant adaptée.

d) Biais résultant des données manquantes :

Pour 5 études, le risque de biais liée aux données manquantes est jugé élevé. Il s'agit des études de Göbel 1994, Pfaffenrath 1994, Brendtsen 1996, Diamond 1998 et Ashina et Brendtsen 2004. Toutes ces études présentent plus de 5% de données manquantes. L'étude de Pfaffenrath utilise la méthode dite "de la dernière visite rapportée" pour la gestion des données manquantes, mais celle-ci est considérée comme étant inadéquate. Hormis pour l'étude de Ashina et Brendtsen 2004, où aucune indication n'est fournie, toutes les autres études ont des données manquantes jugées en lien la vraie valeur du résultat.

L'étude de Brendtsen 2000 est jugée à risque modéré. Bien que cette étude présente 32.5% de données manquantes, celles-ci ne sont a priori pas dépendantes de la vraie valeur du résultat. L'étude de Lance 1964 est jugée à faible risque de biais. En effet, il n'y a pas d'informations données concernant le pourcentage de données manquantes cependant une analyse appropriée en "ITT" a été utilisée avec une minimisation du risque de biais par la règle : chaque donnée manquante pour une molécule est remplacée par "absence d'amélioration".

e) Biais résultant de la mesure du critère de jugement :

Pour 2 études, le risque de biais dans ce domaine est jugé faible : Lance 1964 et Pfaffenrath 1994.

Pour les toutes les études, la méthode de mesure des critères de jugements a été jugée comme étant adaptée. Par ailleurs, les critères de jugements de ces 7 études sont subjectifs et l'évaluateur est tantôt le patient tantôt le patient et le médecin évaluateur. Le maintien du double aveugle est donc indispensable pour garantir l'impartialité de l'évaluation. Or le double aveugle est maintenu dans uniquement 2 études : Lance 1964 et Pfaffenrath 1994, cela explique qu'elles soient à risque faible.

Pour les 5 autres études, nous avons jugé que le double aveugle n'était pas maintenu. Elles sont donc jugées comme étant à risque élevé (Göbel 1994, Brendtsen 1996, Diamond 1998, Brendtsen 2000 et Ashina et Brendtsen 2004).

f) Biais résultant de la sélection des résultats rapportés :

Le risque de biais dans ce domaine est jugé modéré pour 6 études : Lance 1964, Göbel 1994, Pfaffenrath 1994, Brendtsen 1996, Brendtsen 2000, Ashina et Brendtsen 2004.

L'étude de Diamond 1998 est quant à elle jugée à risque élevé.

Pour toutes les études, aucune information n'est fournie concernant l'enregistrement préalable à l'essai du protocole dans un registre ou une base de données. De ce fait nous sommes incapables d'affirmer que les analyses ont été strictement réalisées conformément à un protocole pré-spécifié. Pour toutes les études, les critères de jugement ont été évalués d'une seule manière et tous les résultats étaient présentés.

Pour les six études à risque modéré, il en est de même concernant les analyses. Pour chacune de ces études, une seule analyse est réalisée et présentée.

Pour l'étude de Diamond 1998, deux analyses sont réalisées (une en ITT et une en ITT_m) mais une seule est présentée (ITT). De ce fait cette étude est jugée comme étant à risque élevé.

g) Risque lié à la multiplicité des tests statistiques :

Trois études sur sept ont des critères de jugement principaux multiples mais toutes utilisent une méthode d'ajustement du risque de première espèce et leurs résultats sont donc également

considérés comme “ confirmatoires” (Pfaffenrath 1994, Brendtsen 2000, Ashina et Brendtsen 2004).

L'étude de Pfaffenrath 1994 et de Brendtsen 1986 précisent avoir utilisé la méthode de Bonferroni pour l'ajustement du risque alpha. Les études de Brendtsen 2000 et Ashina et Brendtsen 2004 ne précisent pas l'utilisation d'une méthode d'ajustement, mais compte tenu du fait qu'elles déclarent utiliser la même méthodologie que l'étude de 1996, nous avons supposé qu'une méthode de Bonferroni avait probablement été appliquée.

L'étude de Diamond 1998 évalue plusieurs critères de jugements sans préciser le caractère principal ou secondaire de ceux-ci. Pour cette étude aucune méthode de minimisation du risque alpha n'a été déclarée.

3. Analyse de sensibilité :

Dans notre méthodologie, nous avons statué sur les items suivants :

- La levée du double aveugle en cas d'absence d'information sur les effets indésirables spécifiques de l'amitriptyline
- Un seuil de données manquante et d'acceptabilité de perdus de vue analyse en ITTm à 5%
- Une réponse “ absence d'information “ si les auteurs ne déclarent pas avoir enregistrés leur protocole au préalable

Compte tenu du fait que ces décisions impactent fortement l'évaluation du risque de biais de chaque catégorie, il a été convenu, en accord avec Pr Boussageon, de réaliser une étude de sensibilité méthodologique. Ainsi, pour chaque article scientifique analysé, nous avons appliqué un des items suivants pour voir si ce changement peut impacter positivement le niveau de risque global.

- item 1 : Pour la non mention des effets indésirables une estimation d'absence d'impact sur le double aveugle
- item 2 : Pour les données manquantes et le seuil d'acceptabilité de perdu de vue d'ITTm à 10 et si besoin 15%
- item 3 : Une réponse “ probablement oui “ si les auteurs ne fournissent aucune information sur l'enregistrement du protocole

Par la suite, nous avons décidé de réaliser une analyse de sensibilité en appliquant les 3 items afin de maximiser les chances d'obtenir des études avec un risque acceptable.

A - Lance 1964

En appliquant l'item 3 uniquement : le domaine 5 passe à faible risque.

Pas de modification des domaines en appliquant l'item 1 et 2.

Impact sur le risque global : dans tous les cas risque reste élevé car domaine 1 et 2 classé “
Quelques inquiétudes”

B – Göbel 1994

En appliquant l'item 3 uniquement : le domaine 5 passe à faible risque.

Pas de modification des domaines en appliquant l'item 1 et 2.

Impact sur le risque global : dans tous les cas risque reste élevé car domaines 2, 3 et 4 à risque élevé.

C - Pfaffenrath 1994

En appliquant l'item 3 uniquement : le domaine 5 passe à faible risque.

Pas de modification des domaines en appliquant l'item 1 et 2.

Impact sur le risque global : dans tous les cas risque reste élevé car domaine 3 à risque élevé

D - Brendtsen 1996

En appliquant l'item 2 avec un seuil à 15% : le domaine 2 et 3 passent en risque faible

En appliquant l'item 3 uniquement : le domaine 5 passe à faible risque.

Impact sur le risque global : dans tous les cas risque reste élevé car domaine 4 à risque élevé

E - Diamond 1998

Pas de modification des domaines en appliquant les 3 items

Impact sur le risque global : dans tous les cas risque reste élevé

F - Brendtsen 2000

En appliquant l'item 3 uniquement : le domaine 5 passe à faible risque.

Pas de modification des domaines en appliquant l'item 1 et 2.

Impact sur le risque global : dans tous les cas risque reste élevé car domaine 2 et 4 à risque élevé

G – Ashina et Brendtsen 2004

En appliquant l'item 2 avec un seuil à 15% : le domaine 2 et 3 passent en risque faible

En appliquant l'item 3 uniquement : le domaine 5 passe à faible risque.

Impact sur le risque global : dans tous les cas risque reste élevé car domaine 4 à risque élevé

Malgré notre analyse de sensibilité, réalisée en appliquant les 3 items à chaque fois pour chaque étude, nous constatons que toutes nos études restent à risque élevé de biais.

Cela s'explique par la levée du double aveugle pour certaines études et une proportion très importante de données manquantes (supérieures à 15%) pour d'autres.

Nom l'auteur principal et année de publication	Biais résultant du processus de randomisation	Biais dû à l'effet de période et à l'effet carry over	Biais dus aux écarts aux interventions prévues	Biais résultant des données manquantes	Biais résultant de la mesure du critère de jugement	Biais résultant de la sélection des résultats rapportés	Biais dus à l'inflation du Risque alpha	Evaluation globale
Lance et Curran 1964								Risque élevé
Gobel 1994								Risque élevé
Pfaffenrath 1994								Risque élevé
Brendtsen 1996								Risque élevé
Diamond 1998								Risque élevé
Brendtsen 2000								Risque élevé
Ashina et Brendtsen 2004								Risque élevé

Vert : risque faible ; Orange : risque modéré, Rouge : risque élevé



Figure 2 : Evolution du pourcentage d'études présentant un risque de biais estimé « faible » ou présentant « quelques inquiétudes » ou « élevé », par domaine de biais, selon les 2 interprétations des items du RoB tool 2 (initiale et révisée)

L'analyse de sensibilité méthodologique a été réalisée sur les 7 études incluses dans la méta-analyse. Elle montre que, même après utilisation de l'outil RoB tool 2 avec l'interprétation révisée de ses items, l'évaluation du risque global de biais des études incluses est inchangée. Toutes les études restent avec un risque global de biais jugé élevé. Les changements d'interprétation impactent faiblement les domaines de biais : déviations par rapports aux interventions prévues, et données manquantes (Figure 2).

4. Résultats de chaque étude :

Toutes les études évaluent l'efficacité de l'amitriptyline dans le traitement des céphalées selon la fréquence et/ou la durée et/ou l'intensité des douleurs, les évaluant de manière isolée et/ou en les associant formant ainsi un critère composite.

a) **Lance 1964**

L'étude de Lance 1964 réalise deux essais cliniques distincts. Les deux essais sont des essais en crossover.

Le premier comporte 280 patients randomisés dans 12 bras d'essai : Les participants ont reçu :

- soit des sédatifs (amylobarbitone),
- soit des Vasodilatateurs,
- soit la combinaison d'un sédatif et d'un médicament vasodilatateur,
- soit du "Ballergal " un mélange de phénobarbital, ergotamine et belladone,
- soit du Méthysergide un dérivé d'acide lysergique,
- soit du citrate d'orphénadrine,
- soit du chlordiazepoxide,
- soit du Diazépam,
- soit de l'Amitriptyline
- soit de l'Imipramine,
- soit du "Parstelin" un mélange de tranlycypromine et trifluopérazine,
- soit un Placebo

La durée de l'essai est de 12 mois sans période de wash out ni de période initiale de pré-traitement. L'évaluation de l'amélioration de la douleur selon la fréquence ou l'intensité ("pratiquement la disparition des céphalées", "plus de 50% d'amélioration" ou "pas de modification ") est réalisée toutes les 4 semaines. Lorsque le traitement est jugé efficace le patient poursuit ce même traitement, s'il ne l'est pas il est alors randomisé dans un nouveau bras d'intervention.

Au total 104 patients ont été traités par l'amitriptyline de 10 à 25 mg trois fois par jour au cours de l'essai et 49 par placebo. L'analyse réalisée est une "ITT" avec 7 perdus de vue dans le

groupe amitriptyline et 4 dans le groupe placebo. L'amélioration des symptômes est observée de manière significative :

22 patients ont déclaré la disparition des céphalées dans le groupe Amitriptyline versus 2 dans le groupe placebo.

36 patients ont déclaré avoir une amélioration d'au moins 50% des symptômes dans le groupe amitriptyline versus 9 dans le groupe placebo.

Au total 55% des patients ont déclaré une amélioration totale ou supérieure à 50% contre 22% pour le groupe placebo. Le seuil de significativité n'est pas fourni.

Le deuxième essai est un essai qui dure deux mois, toujours en crossover évaluant exclusivement l'amitriptyline contre un placebo sans période de wash out. Au total 27 patients ont été inclus. L'évaluation est réalisée à la fin du premier mois de traitement. Le critère de jugement principal est "amélioration ou non des céphalées". L'analyse réalisée est une "ITT" avec 1ss perdus de vue dans le groupe amitriptyline et 2 dans le groupe placebo.

L'amélioration des symptômes est observée de manière significative :

9 patients ont déclaré la disparition des céphalées dans le groupe amitriptyline versus 1 dans le groupe placebo,

15 patients ont déclaré avoir une amélioration d'au moins 50% des symptômes dans le groupe amitriptyline versus 5 dans le groupe placebo.

Au total 55% des patients ont déclaré une amélioration totale ou supérieure à 67% contre 30% pour le groupe placebo. Le seuil de significativité n'est pas fourni. *"This result is highly significant by the modified test of the fourfold contingency table (Fisher 1961) and by the exact method using the binomial distribution"*.

Concernant les effets secondaires, 29 patients (28%) se sont plaints d'effets secondaires suite à la prise d'amitriptyline particulièrement la somnolence, les tremblements et la sécheresse buccale. Parmi ceux-ci 7 ont abandonné l'étude.

Pour le groupe placebo 4 patients ont abandonné l'étude pour un urticaire, une sensation de dépression, douleur abdominale et nausée et vertiges. Il est important de noter que dans cette étude l'amitriptyline n'est pas la seule molécule pouvant causer une somnolence ou une sécheresse buccale, ce qui permet de préserver l'aveugle.

b) Göbel 1994

L'étude de Göbel 1994 évalue l'efficacité de l'amitriptyline versus un placebo dans un essai en groupes parallèles d'une durée de 6 semaines. Au total 78 patients ont été inclus, 24 dans le groupe amitriptyline et 29 dans le groupe placebo. Parmi les 78 patients, 25 patients ont été exclus dont 4 perdus de vue pour effets secondaires et 21 pour violation du protocole.

L'amitriptyline est administrée sous la forme d'hydrochlorure d'amitriptyline en libération prolongé à la dose de 25 mg, le dosage était progressivement augmenté : 1 gélule la première semaine, 2 gélules (50mg) la seconde et 3 gélules (75mg) à partir de la 3^{ème} semaine.

L'évaluation est réalisée toutes les semaines. Le critère de jugement principal est la modification de l'activité EMG des muscles péricrâniens, dans la suppression extéroceptive de l'activité du muscle temporal, du CNV et de la sensibilité à la douleur lors d'un traitement par amitriptyline pour des céphalées de tension. Pour cela, la durée journalière moyenne des céphalées par semaine est également évaluée. L'analyse réalisée est en "per protocole".

Le résultat montre une **amélioration significative de la durée moyenne des céphalées par semaine** pour le groupe amitriptyline versus placebo à la sixième semaine de traitement seulement avec pour le groupe amitriptyline 7,9 heures (+/- 6.0) et pour le groupe placebo 9,6 heures (+/- 6.0) avec $p=0.007$.

Concernant les effets secondaires, les auteurs indiquent ne pas avoir observé de différence entre les groupes concernant les effets secondaires globaux, cependant **une différence significative** a été observée pour les effets secondaires spécifiques. En effet 54% des patients du groupe amitriptyline ont signalé une sécheresse buccale contre 17 % dans le groupe placebo ($p<0,05$) et 62% des patients du groupe amitriptyline ont noté une somnolence contre 27% dans le groupe placebo ($p<0,05$). Concernant les abandons, 3 patients ont abandonné l'étude dans le groupe Placebo versus 1 dans le groupe amitriptyline.

c) Pfaffenrath 1994

L'étude de Pfaffenrath 1994 est un essai en groupe parallèle à 3 bras d'intervention comparant l'efficacité de l'amitriptyline à de l'amitriptylinoxide et à un placebo sur une durée de 24 semaines.

Au total 197 patients ont été inclus, avec 66 patients dans le groupe amitriptylinoxide, 67 dans le groupe amitriptyline et 64 dans le groupe placebo. Par ailleurs 62 patients ont été perdus de vue.

La dose initiale d'amitriptyline est de 25 mg, une augmentation progressive est réalisée de telle sorte qu'à partir de 9 semaines les participants passent à 2 comprimés de 25mg et à partir de la 13^{ème} semaine à 3 comprimés.

Plusieurs critères de jugement principaux sont mesurés :

- La fréquence des céphalées (nb jours/mois),
- La durée des céphalées (nb d'heures de céphalées/jours multiplié par mois)
- L'intensité des céphalées (VAS de 0 à 8).
- Enregistrement des effets indésirables avec le questionnaire FSUCL

A partir de ces mesures les auteurs ont estimé :

- Le score fréquence des céphalées x durée des céphalées
- L'amélioration des patients après 12 semaines selon : durée multipliée par la fréquence, durée multipliée par la Fréquence et l'Intensité, l'Intensité seule.
- La réduction d'au moins 50 % du produit durée multiplié par la fréquence et au moins une réduction de 50 % dans l'intensité des céphalées."

L'évaluation est réalisée toutes les 4 semaines.

L'analyse réalisée est en "intention de traiter".

Les résultats retrouvent une [absence d'amélioration significative pour l'amitriptyline](#) versus un placebo pour tous ces critères de jugement.

Concernant la durée des céphalées : dans le groupe amitriptyline on enregistre 7 heures en moyenne par jour et par mois $\pm 7h$ et dans le groupe placebo 7 heures $\pm 6h$.

Concernant l'intensité des céphalées : dans le groupe amitriptyline on enregistre une EVA à 2.8 ± 2.0 tandis que dans le groupe placebo 1.7 ± 2.0 .

Concernant la fréquence des céphalées : dans le groupe amitriptyline on enregistre 15 jours par mois ± 10 et dans le groupe placebo 16 jours par mois ± 9 .

Le score fréquence multiplié par la durée (fréquence x durée) retrouve pour le groupe amitriptyline une moyenne de 207 avec un écart type à ± 194 et pour le groupe placebo 208 avec un écart type à ± 171 .

Après 12 semaines de patients, pour le score durée multiplié par la fréquence (durée x fréquence), 41.8 % des patients sous amitriptyline se disent améliorés tandis qu'ils sont 28.1 % dans le groupe placebo. Pour la durée multipliée par la fréquence multipliée par l'intensité (durée x fréquence x intensité) ils sont 28.4 % dans le groupe amitriptyline et 20.3 % dans le groupe placebo. Pour l'intensité, ils sont 32.8 % dans le groupe amitriptyline et 28.1% dans le groupe placebo.

Concernant les effets secondaires, l'enregistrement des effets indésirables via le questionnaire FSUCL n'a pas retrouvé de différence significative entre les trois groupes. Concernant les abandons pour cause d'effets indésirables ils ont concerné 3% des patients dans le groupe amitriptylinoxyde, 17,9 % dans le groupe Amitriptyline et 10,9% dans le groupe Placebo.

d) Brendtsen 1996

L'étude de Brendtsen 1996 est un essai en crossover à 3 bras d'intervention comparant l'efficacité de l'amitriptyline au citalopram et à un placebo. La durée totale de l'essai est de 32 semaines avec une période initiale "run in period" de 4 semaines, deux périodes de "wash out" de 2 Semaines et 3 périodes de traitement de 8 semaines.

L'amitriptyline est testée à la dose de 25mg par jour initialement augmentée à 50 mg la deuxième semaine puis à 75 mg par jour à partir de la 3^{ème} semaine. Le citalopram est testé à la dose de 20mg.

Au total 40 patients ont été inclus, parmi lesquels il y a eu 6 patients perdus de vue. Le critère de jugement principal est l'aire sous la courbe des céphalées calculée comme la somme des enregistrements quotidiens de la durée de la céphalée x l'intensité de la céphalée. Le résultat est mesuré les 4 dernières semaines de chaque période de traitement. L'analyse est réalisée en "ITTm".

En critères de jugement secondaire, les auteurs ont mesuré la durée des céphalées en heures au cours des quatre dernières semaines de chaque période de traitement, l'intensité moyenne des céphalées par jour de céphalées sur une EVA de 0 à 10, la fréquence des céphalées en nombre de jours au cours des quatre dernières semaines de chaque période de traitement

Les résultats retrouvent une **amélioration significative** pour l'amitriptyline versus placebo :

- Aire sous la courbe (AUC) significativement plus basse pour le groupe amitriptyline à 616 avec un écart type de 129 versus 877 avec un écart type de 171. Réduction de 30 % de la courbe AUC ($p = 0.002$) pour l'amitriptyline en comparaison au placebo.
- Diminution significative de la durée avec pour l'amitriptyline une durée mesurée par mois à 151 heures avec un écart type à 24 heures et pour le placebo une durée de 184 heures avec un écart type de 27 heures ($p = 0.01$).
- Diminution significative de la fréquence des céphalées avec pour l'amitriptyline 18.6 jours avec un écart type de 1.6 jours et pour le groupe placebo 21.7 jours avec un écart type de 1.3 jours ($p = 0.01$).

- Diminution non significative de l'intensité des céphalées avec pour le groupe amitriptyline une EVA moyenne à 3.8 avec un écart type de 0,3 versus 3.9 pour le groupe placebo avec un écart type de 0,3.

Concernant les effets secondaires, les auteurs retrouvent une occurrence **significativement plus élevée** dans le groupe amitriptyline pour les effets indésirables généraux signalés par 33 patients dans groupes amitriptyline versus 15 dans le groupe placebo ($p < 0.001$). Il en est de même pour les effets secondaires spécifiques avec une sécheresse buccale chez 30 patients dans le groupe amitriptyline versus 3 dans le groupe placebo ($p < 0,001$) et la survenue d'une somnolence chez 21 patients dans le groupe amitriptyline versus 6 dans le groupe placebo ($p < 0,001$). Il y a eu par ailleurs 6 patients perdus de vue sans plus de détails apportés.

e) Diamond 1998

L'étude de Diamond 1998 évalue l'efficacité de l'amitriptyline à deux dosages différents 10 mg et 25 mg versus un placebo dans un essai en groupes parallèles à 3 bras d'intervention pour une durée de 4 semaines. Au total 90 patients ont été inclus, 30 dans le groupe amitriptyline 25mg, 30 dans le groupe amitriptyline 10 mg et 30 dans le groupe placebo. Parmi les 90 patients, 30 patients ont été exclus dans l'analyse des résultats pour céphalées. 5 ont été abandonnée durant la phase initiale de l'essai.

L'évaluation est réalisée toutes les semaines. Il n'y a pas d'indication concernant le caractère principal ou secondaire des critères de jugements. Les céphalées sont évaluées sur une échelle de 1 à 4 avec 1= aucun, 2= léger, 3= modéré, 4=sévère. L'analyse réalisée pour les résultats de la céphalée est une ITTm.

Le résultat est significatif pour l'amitriptyline 10 mg versus placebo dans le traitement des céphalées à 1 semaine uniquement avec une moyenne à 1.5 pour le groupe amitriptyline versus 0.6 pour le groupe placebo ($p < 0,05$).

Concernant les effets secondaires, ils ont concerné 8 patients dans le groupe amitriptyline 10 mg, 8 dans le groupe amitriptyline 25 mg et seulement 1 patient dans le groupe placebo. **Cette différence est statistiquement significative** ($p < 0,05$). La sécheresse buccale a concerné 3 patients dans le groupe amitriptyline 10 mg, 5 patients dans le groupe amitriptyline 25 mg et aucun dans le groupe placebo. La somnolence a concerné 3 patients dans le groupe amitriptyline 10 mg, 5 patients dans le groupe amitriptyline 25 mg et aucun dans le groupe placebo. Pour les effets secondaires il n'est pas précisé si les différences sont statistiquement

significatives. Il y a eu 9 patients perdus de vue dans le groupe amitriptyline versus 5 dans le groupe Placebo.

f) Brendtsen 2000

L'étude de Brendtsen 2000 est un essai en crossover à 3 bras d'intervention comparant l'efficacité de l'amitriptyline au citalopram et à un placebo. La durée totale de l'essai est de 32 semaines avec une période initiale "run in period" de 4 semaines, deux périodes de "wash out" de 2 Semaines et 3 périodes de traitement de 8 semaines.

L'amitriptyline est testée à la dose de 25mg par jour initialement augmentée à 50 mg la deuxième semaine puis à 75 mg par jour à partir de la 3^{ème} semaine. Le citalopram est testé à la dose de 20 mg.

Au total 40 patients ont été inclus, parmi lesquels il y a eu 6 patients perdus de vue. Les auteurs évaluent en critères de jugement secondaire : l'aire sous la courbe des céphalées calculée comme la somme des enregistrements quotidiens de la durée de la céphalée x l'intensité de la céphalée et l'intensité moyenne des céphalées par jour de céphalées sur une EVA de 0 à 100 mm.

Le résultat est mesuré les 4 dernières semaines de chaque période de traitement. L'analyse est réalisée en "ITTm".

Résultats significatifs pour l'amitriptyline versus placebo pour :

- La courbe AUC avec une réduction de 30% de l'aire sous la courbe pour le groupe amitriptyline versus placebo, résultat qualifié "hautement significatif " sans autre précision.
- L'intensité des céphalées avec pour le groupe amitriptyline une EVA médiane à 10 avec un quartile de 0 à 20 versus 16 pour le groupe placebo avec un quartile de 0 à 37.

Par ailleurs, les auteurs ne donnent aucune information concernant l'occurrence d'effets secondaires. Les auteurs précisent qu'il y a eu 7 patients perdus de vue, mais ces abandons n'étaient pas liés aux effets indésirables.

g) Ashina et Brendtsen 2004

L'étude d'Ashina et Brendtsen 2004, tout comme les autres études de l'auteur Brendtsen, est également un essai en crossover à 3 bras d'intervention, comparant l'efficacité de l'amitriptyline au citalopram et à un placebo. La durée totale de l'essai est de 32 semaines avec une période initiale "run in period" de 4 semaines, deux périodes de "wash out" de 2 semaines et 3 périodes de traitement de 8 semaines.

L'amitriptyline est testée à la dose de 25 mg par jour initialement augmentée à 50 mg la deuxième semaine puis à 75 mg par jour à partir de la 3^{ème} semaine. Le citalopram est testé à la dose de 20 mg.

Au total 40 patients ont été inclus, parmi lesquels il y a eu 6 patients perdus de vue.

Les auteurs évaluent en critère de jugement principal :

- L'aire sous la courbe des céphalées calculée comme la somme des enregistrements quotidiens de la durée de la céphalée x l'intensité de la céphalée
- La durée des céphalées en heures au cours des quatre dernières semaines de chaque période de traitement
- L'intensité moyenne des céphalées par jour de céphalées sur une EVA de 1 à 10.

Le résultat est mesuré les 4 dernières semaines de chaque période de traitement. L'analyse est réalisée en "ITTm".

Résultats significatifs pour l'amitriptyline versus placebo avec diminution significative de l'aire sous la courbe AUC avec pour l'amitriptyline une médiane à 308 avec comme quartile (157 – 715) versus 441 (178-1408) pour le placebo avec $p=0.001$.

Par ailleurs, les auteurs ne donnent aucune information concernant l'occurrence d'effets secondaires.

Les auteurs ne donnent pas plus de détails concernant les 6 patients perdus de vue.

5. Synthèse des résultats :

a) Synthèse qualitative

Concernant la durée des céphalées :

3 études pour 252 patients au total, Amitriptyline : 125 patients, Placebo : 127 patients.

- **Diminution significative** de la durée moyenne des céphalées dans l'étude Göbel 1994 en critère de jugement principal : uniquement pour la sixième semaine avec pour le groupe Amitriptyline $7,9 \pm 6$ heures par semaine contre $9,6 \pm 6$ heures par semaine pour le groupe placebo ($p = 0.007$).
- **Diminution significative** de la durée dans l'étude de Brendtsen 1996 en critère de jugement secondaire avec pour l'amitriptyline une durée mesurée à 151 heures par 4 semaines avec un écart type à 24 heures et pour le placebo une durée de 184 heures avec un écart type de 27 heures ($p = 0.01$).
- **Résultats non significatifs** dans l'étude Pfaffenrath 1994 en critère de jugement principal : dans le groupe amitriptyline on enregistre 7 heures en moyenne par jour et par quatre semaines $\pm 7h$ et 7 ± 6 heures dans le groupe placebo.

Concernant l'intensité des céphalées :

3 études pour un total de 267 patients, Amitriptyline : 135 patients, Placebo 132 patients.

- **Diminution significative** de l'intensité des céphalées dans l'étude Brendtsen 2000 en critère de jugement secondaire avec pour le groupe amitriptyline la médiane de l'EVA à 10 avec un quartile de 0 à 20 versus 16 pour le groupe placebo avec un quartile de 0 à 37.
- **Résultats non significatifs** dans l'étude Pfaffenrath 1994 en critère de jugement principal : dans le groupe amitriptyline on enregistre une EVA à 2.8 ± 2.0 tandis que dans le groupe placebo 1.7 ± 2.0 . Ce résultat est donc en faveur du placebo.
- **Résultats non significatifs** dans l'étude Brendtsen 1996 en critère de jugement secondaire : pour le groupe amitriptyline une EVA moyenne à 3.8 avec un écart type de 0,3 versus 3.9 pour le groupe placebo avec un écart type de 0,3.

Concernant la fréquence des céphalées :

2 études pour un total de 199 patients, Amitriptyline: 101 patients, Placebo :98 patients

- **Diminution significative** de la fréquence des céphalées dans l'étude de Brendtsen 1996 en critère de jugement secondaire avec pour l'amitriptyline 18.6 jours avec un écart type de 1.6 jours et pour le groupe placebo 21.7 jours avec un écart type de 1.3 jours ($p = 0.01$).
- **Résultats non significatifs** dans l'étude Pfaffenrath 1994 en **critère de jugement principal** : dans le groupe amitriptyline on enregistre 15 ± 10 jours par mois contre 16 ± 9 jours par mois dans le groupe placebo.

Concernant l'amélioration des céphalées de manière indifférenciée :

L'étude de Diamond 1998 retrouve un **résultat significatif** avec pour l'amitriptyline, à 1 semaine uniquement, une moyenne à 1.5 pour le groupe versus 0.6 pour le groupe placebo ($p < 0,05$) sur une échelle de 1 à 4 avec 1= aucun, 2= léger, 3= modéré, 4=sévère.

Concernant la mesure composite de l'intensité et/ou la fréquence :

- **Amélioration significative** des céphalées dans les deux études de Lance 1964 en **critère de jugement principal** avec :
 - *Pour le premier essai* : 22 patients sur 104 ont déclaré la disparition des céphalées dans le groupe Amitriptyline versus 2 sur 49 dans le groupe placebo. 36 patients sur 104 ont déclaré avoir une amélioration d'au moins 50% des symptômes dans le groupe amitriptyline versus 9 sur 49 dans le groupe placebo. Au total 55% des patients ont déclaré une amélioration totale ou supérieure à 50% contre 22% pour le groupe placebo. Seuils de significativité non fournis.
 - *Pour le second essai* : Sur 27 patients, 9 patients ont déclaré la disparition des céphalées dans le groupe amitriptyline versus 1 dans le groupe placebo, 15 patients ont déclaré avoir une amélioration d'au moins 50% des symptômes dans le groupe amitriptyline versus 5 dans le groupe placebo. Au total 55% des patients ont déclaré une amélioration totale ou supérieure à 67% contre 30% pour le groupe placebo. Le seuil de significativité n'est pas fourni.

Concernant la mesure composite de l'intensité et de la durée :

- Amélioration significative des moyennes des sommes des enregistrements quotidiens de la durée multipliée par l'intensité, présentées sous forme d'aire sous la courbe AUC dans les études de Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004.

- Brendtsen 1996 en **critère de jugement principal** : Aire sous la courbe (AUC) **significativement plus basse** pour le groupe amitriptyline à 616 avec un écart type de 129 versus 877 avec un écart type de 171. Réduction de 30 % de la courbe AUC ($p=0.002$) pour l'amitriptyline en comparaison au placebo.
- Brendtsen 2000 en critère de jugement secondaire : réduction de 30% de l'aire sous la courbe pour le groupe amitriptyline versus placebo, résultat qualifié **“hautement significatif”** sans autre précision.
- Ashina et Brendtsen 2004 en critère de jugement secondaire : **diminution significative** de l'aire sous la courbe AUC avec pour l'amitriptyline une médiane à 308 avec comme quartile (157 – 715) versus 441 (178-1408) pour le placebo avec $p=0.001$.

Toutes les autres associations de mesure entre fréquence, durée et intensité retrouvent des **résultats non significatifs** dans l'étude Pfaffenrath 1994 en critères de jugement secondaires : **Pour le score fréquence x durée** : le groupe amitriptyline a un score moyen de 207 ± 194 et le groupe placebo 208 ± 171 .

Après 12 semaines :

- **Pour le score durée x fréquence** : 41.8 % des patients sous amitriptyline se disent améliorés tandis qu'ils sont 28.1 % dans le groupe placebo.
- **Pour la durée x fréquence x intensité** : 28.4 % dans le groupe amitriptyline se disent améliorés contre 20.3 % dans le groupe placebo.
- **Pour l'intensité** : 32.8 % dans le groupe amitriptyline se disent améliorés contre 28.1% dans le groupe placebo.

Concernant la mesure des effets secondaires :

Deux études retrouvent une différence significative concernant les effets secondaires généraux :

L'étude de Brendtsen 1996 retrouve une occurrence **significativement plus élevée** dans le groupe amitriptyline avec 33 patients concernés dans groupes amitriptyline versus 15 dans le groupe placebo ($p < 0.001$).

L'étude de Diamond 1998 retrouve également une **différence statistiquement significative** ($p < 0,05$) avec 8 patients concernés dans le groupe amitriptyline 10 mg, 8 dans le groupe amitriptyline 25 mg et seulement 1 patient dans le groupe placebo.

L'étude de Lance 1964 retrouve une différence concernant les effets secondaires généraux et spécifiques sans précision sur la significativité statistique : 29 patients (28%) se sont plaints d'effets secondaires suite à la prise d'amitriptyline particulièrement la somnolence, les tremblements et la sécheresse buccale. Parmi ceux-ci 7 ont abandonné l'étude. Pour le groupe placebo 4 patients ont abandonné l'étude pour un urticaire, une sensation de dépression, douleur abdominale et nausée et vertiges.

L'étude de Göbel 1994 indique ne pas avoir observé de différence entre les groupes concernant les effets secondaires généraux.

L'étude de Pfaffenrath 1994 n'a pas trouvé de différence significative entre les trois groupes pour les effets indésirables via le questionnaire FSUCL.

Deux études retrouvent une différence significative concernant les effets secondaires spécifiques :

L'étude de Göbel 1994 retrouve [une différence significative](#) avec 54% des patients du groupe amitriptyline présentant une sécheresse buccale contre 17 % dans le groupe placebo ($p<0,05$) et 62% des patients du groupe amitriptyline ayant signalé une somnolence contre 27% dans le groupe placebo ($p<0,05$).

L'étude de Brendtsen 1996 retrouve également [une différence statistiquement significative](#) avec une sécheresse buccale chez 30 patients dans le groupe amitriptyline versus 3 dans le groupe placebo ($p<0,001$) et la survenue d'une somnolence chez 21 patients dans le groupe amitriptyline versus 6 dans le groupe placebo ($p<0,001$)

L'étude de Diamond 1998 montre une différence concernant les effets indésirables spécifiques sans préciser la significativité statistique : sécheresse buccale chez 3 patients dans le groupe amitriptyline 10 mg, 5 patients dans le groupe amitriptyline 25 mg et aucun dans le groupe placebo. La somnolence a concerné 3 patients dans le groupe amitriptyline 10 mg, 5 patients dans le groupe amitriptyline 25 mg et aucun dans le

L'étude de Lance 1964 retrouve également une différence sans précision sur la significativité statistique.

Les études de Brendtsen 2000 et Ashina et Brendtsen 2004 ne rapportent pas d'informations concernant l'occurrence des effets secondaires.

Résultats à caractère confirmatoire :

[Les études suivantes apportent des résultats significatifs et sont considérées comme étant “confirmatoire” pour ces résultats :](#)

- la durée des céphalées dans l'étude de Göbel 1994
- la mesure composite de l'intensité et/ou la fréquence dans l'étude de Lance 1964
- la mesure composite de l'intensité et de la durée (courbe AUC) dans Brendtsen 1996

L'étude de Pfaffenrath apporte des résultats non significatifs et est considérée comme étant “confirmatoire” pour ces résultats car elle utilise une méthode d'ajustement du risque alpha :

- Durée, fréquence et intensité des céphalées, ainsi que pour l'enregistrement des effets indésirables.

b) Synthèse quantitative

La réalisation d'une méta analyse n'a pas été possible car plusieurs problèmes se sont posés :

- Plusieurs auteurs utilisent des scores composites propres à leur étude
- Certains auteurs n'ont pas fourni le détail des résultats pour chaque effectif de patients. Pour pallier ce problème nous avons tenté de joindre les auteurs malheureusement soit ils n'ont pas donné de réponse, soit ils n'avaient plus les données en leur possession.
- Une méta analyse pour la durée des céphalées n'a pas été possible car les trois auteurs n'utilisent pas les mêmes unités de mesures (en heure par semaine pour Göbel 1994, heure par jour par mois pour Pfaffenrath 1994 et en heure par mois pour Brendtsen 1996) et la conversion des résultats n'est pas possible.
- Une méta analyse pour l'intensité des céphalées n'a pas été possible car l'étude de Brendtsen 2000 fournit une médiane de l'EVA contrairement aux deux autres études qui fournissent une moyenne.
- Une méta analyse pour la fréquence des céphalées n'a pas non plus été réalisée car nous ne disposons que de deux études.

6. Evaluation du niveau de preuve global selon REB et GRADE :

❖ Selon la méthode REB (13)

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2 et les données extraites ne nous ont pas permis de réaliser de méta analyse de sensibilité. Pour le critère de jugement “ durée” deux études sur trois sont significativement en faveur de l'amitriptyline. Pour les autres critères, la majorité des études ne sont pas significatives. Par ailleurs nous ne sommes pas en mesure d'affirmer qu'il n'existe pas de biais de publication.

L'évaluation de la douleur selon la durée, l'intensité et la fréquence selon la méthode REB conclue à un manque de preuve concernant l'efficacité de l'amitriptyline dans les céphalées de tension chroniques.

❖ Selon la méthode GRADE (14)

L'outil GRADE permet de rendre un jugement sur la qualité des résultats et donc sur le niveau de preuve.

L'outil GRADE n'évalue pas un essai clinique mais un résultat. Plusieurs analyses différentes peuvent donc être réalisées pour un même essai clinique.

La note initiale attribuée à un essai contrôlé randomisé est une note de base élevée. Cette note de référence peut ensuite être abaissée ou améliorée lors de l'évaluation des 8 critères ci-dessous :

Raisons de dégrader la qualité des preuves :

1. Risque de biais
2. Incohérence
3. Le caractère indirect des résultats
4. Imprécision
5. Biais de publication

Raisons d'améliorer la qualité des preuves :

6. Grande ampleur de l'effet
7. Réponse à la dose

8. L'effet de tous les facteurs de confusion plausibles serait de réduire l'effet (lorsqu'un effet est observé) ou suggèrent un effet parasite (lorsqu'aucun effet n'est observé)

a) Niveau de preuve pour “ la durée des céphalées “

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

1. Risque de biais : Les études de Göbel 1994, Pfaffenrath 1994 et Brendtsen 1996 évaluent la durée des céphalées. Ces trois études sont à risque élevé de biais. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, deux études : Göbel 1994 et Brendtsen 1996 présentent un risque jugé élevé pour 3 domaines du RoB 2 : risque lié aux écarts aux interventions prévues, risque lié aux données manquantes et risque lié à l'évaluation des résultats. L'étude de Pfaffenrath 1994 a un risque de biais élevé pour le domaine lié aux données manquantes. De ce fait nous décidons de retirer deux points.
2. Incohérence : Les 3 études suggèrent toutes une différence significative en faveur de l'amitriptyline, de manière significative pour Göbel 1994 et Brendtsen 1996 et de manière non significative pour Pfaffenrath 1994. Les autres critères évaluant l'hétérogénéité ne peuvent pas être appliqués dans ce cas présent car nous n'avons pas pu réaliser de méta analyse. Cependant les auteurs utilisent des unités de mesure différente rendant l'interprétation des résultats plus approximative. Nous décidons de retirer deux points pour cet item.
3. Le caractère indirect des résultats : Les trois études sont en lien avec l'objectif de notre revue de la littérature. La population étudiée, l'intervention ainsi que le comparateur correspondent également à l'objectif de notre étude. Nous décidons de ce fait de ne pas retirer de point.
4. Imprécision : Les trois études ont chacune moins de 400 patients inclus : 78 pour Göbel 1994, 197 pour Pfaffenrath 1994 et uniquement 40 pour Brendtsen 1996. De ce fait nous décidons de retirer un point.
5. Biais de publication : Les études sont toutes à petits échantillons de ce fait nous pouvons exclure un biais de publication. De ce fait nous retirons également un point.

Nous n'avons pas rajouté de points pour les items “ ampleur de l'effet”, “ effet dose” et “ biais plausibles “.

En prenant en compte ces résultats, nous concluons donc que la preuve de l'efficacité de l'amitriptyline est donc de très faible niveau de preuve pour la réduction de la durée des céphalées.

b) Niveau de preuve pour “ l'intensité des céphalées “

1. Risque de biais : Les études de Pfaffenrath 1994, Brendtsen 1996 et Brendtsen 2000 évaluent l'intensité des céphalées. Ces trois études sont à risque élevé de biais. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, l'étude de Brendtsen 1996 présente un risque jugé élevé pour 3 domaines du RoB 2 (domaine lié aux écarts aux interventions prévues, risque lié aux données manquantes et risque lié à l'évaluation des résultats, L'étude de Brendtsen 2000 a un risque élevé pour deux domaines (domaine lié aux écarts aux interventions prévues et lié à l'évaluation des résultats) et l'étude de Pfaffenrath 1994 a un risque de biais élevé pour le domaine lié aux données manquantes. De ce fait nous décidons de retirer deux points.
2. Incohérence : Les 3 études suggèrent toutes une différence significative en faveur de l'amitriptyline, de manière significative pour Brendtsen 2000 et de manière non significative pour Pfaffenrath 1994 et Brendtsen 1996. Les autres critères évaluant l'hétérogénéité ne peuvent pas être appliqués dans ce cas présent car nous n'avons pas pu réaliser de méta analyse. Cependant les auteurs utilisent des unités de mesure différentes rendant l'interprétation des résultats approximative. Nous décidons de retirer deux points pour cet item.
3. Le caractère indirect des résultats : Les trois études sont en lien avec l'objectif de notre revue de la littérature. La population étudiée, l'intervention ainsi que le comparateur correspondent également à l'objectif de notre étude. Nous décidons de ce fait de ne pas retirer de point.
4. Imprécision : Les trois études ont chacune moins de 400 patients inclus : 78 pour Göbel 1994 et uniquement 40 pour Brendtsen 1996 ainsi que pour Brendtsen 2000. De ce fait nous retirons un point.
5. Biais de publication : Les études sont toutes à petits échantillons de ce fait nous ne pouvons exclure un biais de publication. De ce fait nous retirons également un point.

Nous n'avons pas rajouté de points pour les items " ampleur de l'effet", " effet dose" et " biais plausibles ".

En prenant en compte ces résultats, nous concluons donc que la preuve de l'efficacité de l'amitriptyline est donc de **très faible niveau de preuve pour la réduction de l'intensité des céphalées**.

c) Niveau de preuve pour " la fréquence des céphalées "

1. Risque de biais : Les études de Pfaffenrath 1994, Brendtsen 1996 évaluent la fréquence des céphalées. Ces deux études sont à risque élevé de biais. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, l'étude de Brendtsen 1996 présente un risque jugé élevé pour 3 domaines du RoB 2 (domaine lié aux écarts aux interventions prévues, risque lié aux données manquantes et risque lié à l'évaluation des résultats et L'étude de Pfaffenrath 1994 a un risque de biais élevé pour le domaine lié aux données manquantes. De ce fait nous décidons de retirer deux points
2. Incohérence : Les deux études suggèrent une différence significative en faveur de l'amitriptyline, de manière significative pour Brendtsen 1996 et de manière non significative pour Pfaffenrath 1994. Les autres critères évaluant l'hétérogénéité ne peuvent pas être appliqués dans ce cas présent car nous n'avons pas pu réaliser de méta analyse. Cependant les auteurs utilisent des unités de mesure différentes rendant l'interprétation des résultats approximative. Nous décidons de retirer deux points pour cet item.
3. Le caractère indirect des résultats : Les deux études sont en lien avec l'objectif de notre revue de la littérature. La population étudiée, l'intervention ainsi que le comparateur correspondent également à l'objectif de notre étude. Nous décidons de ce fait de ne pas retirer de point.
4. Imprécision : Les deux études ont chacune moins de 400 patients inclus : 197 pour Pfaffenrath 1994 et uniquement 40 pour Brendtsen 1996. De ce fait nous décidons de retirer un point.
5. Biais de publication : Les études sont toutes à petits échantillons de ce fait nous ne pouvons exclure un biais de publication. De ce fait nous retirons également un point.

Nous n'avons pas rajouté de points pour les items " ampleur de l'effet", " effet dose" et " biais plausibles ".

En prenant en compte ces résultats, nous concluons donc que la preuve de l'efficacité de l'amitriptyline est donc de très faible niveau de preuve pour la réduction de la fréquence des céphalées.

Selon la méthode GRADE, le niveau de preuve est " très faible "concernant l'efficacité de l'amitriptyline dans le traitement des céphalées de tension chroniques en ce qui concerne la durée, l'intensité et la fréquence des céphalées.

V- DISCUSSION

1. Interprétation des résultats :

Notre revue systématique inclut sept essais cliniques randomisés qui étudient l'efficacité de l'amitriptyline dans le traitement des céphalées selon trois modalités : la durée des céphalées, l'intensité et la fréquence des céphalées.

Les sept études sont à haut risque de biais global. Aucune méta-analyse n'a pu être réalisée du fait de l'hétérogénéité des mesures de résultats et/ou de l'absence de données suffisantes fournies par les auteurs pour certaines études.

L'analyse qualitative suggère une différence en faveur de l'amitriptyline dans la diminution de la durée des céphalées, avec deux études montrant une différence significative (Göbel 1994 et Brendtsen 1996) et une étude montrant une différence non significative (Pfaffenrath 1994).

Les études de Göbel 1994 et de Pfaffenrath évaluent la durée des céphalées en tant que critère de jugement principal. Göbel 1994 l'évalue en critère principal unique tandis qu'il fait partie de plusieurs critères de jugement principaux pour Pfaffenrath 1994. Pfaffenrath 1994 utilise une méthode d'ajustement du risque alpha de ce fait ces deux études sont considérées comme confirmatoires pour ce résultat.

Une différence en faveur de l'amitriptyline est également observée dans la diminution de l'intensité des céphalées avec deux études qui suggèrent une différence en faveur de l'amitriptyline (Brendtsen 1996 et 2000), de manière significative pour Brendtsen 2000 et de non significative pour Brendtsen 1996. L'étude de Pfaffenrath 1994 suggère a contrario une différence en faveur du placebo mais qui n'est pas significative. Seule l'étude de Pfaffenrath 1994 est considérée comme confirmatoire pour ce résultat car même si elle possède plusieurs critères de jugement principaux elle utilise une méthode d'ajustement du risque alpha.

Concernant la fréquence des céphalées, deux études suggèrent une différence en faveur de l'amitriptyline mais de manière non significative. Il s'agit des études de Brendtsen 1996 et de Pfaffenrath 1994. Seule l'étude de Pfaffenrath 1994 est considérée comme confirmatoire pour ce résultat car même si elle possède plusieurs critères de jugement principaux elle utilise une méthode d'ajustement du risque alpha.

L'étude de Diamond 1998 évalue l'amélioration globale des céphalées. Le résultat est significatif en faveur de l'amitriptyline uniquement à la fin de la première semaine. Cette étude ne précise pas le caractère principal ou secondaire des critères de jugement choisis. Par ailleurs aucune méthode d'ajustement n'est utilisée. Cette étude est donc considérée comme étant exploratoire.

Les différents scores associant de manière variable la durée, l'intensité et la fréquence ont montré des différences significatives en faveur de l'amitriptyline dans les études de Lance 1964, de Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004. Lance 1964 évalue l'intensité et/ou la fréquence tandis que les études de Brendtsen 1996, Brendtsen 2000 et Ashina et Brendtsen 2004 évaluent la durée x l'intensité. L'étude de Lance 1964 est considérée comme confirmatoire car elle évalue ce score en tant que critère de jugement principal unique. Pour l'étude de Brendtsen 1996 le score utilisé n'est pas le seul critère de jugement principal cependant les auteurs utilisent une méthode d'ajustement du risque alpha. Bien que l'étude de Brendtsen n'évalue pas en priorité l'amitriptyline mais le citalopram, le résultat est également considéré comme confirmatoire.

Les autres associations des mesures réalisées dans l'étude de Pfaffenrath 1994, le score fréquence multipliée par la durée ou l'évaluation du pourcentage de patients améliorée selon les critères durée multipliée par la fréquence, durée multipliée par la fréquence multipliée par l'intensité ou intensité seule, ont montré une différence en faveur de l'amitriptyline mais qui n'est pas significative. Ces résultats sont considérés comme exploratoires.

Concernant les effets indésirables, bien que les résultats ne soient pas toujours significatifs, toutes les études suggèrent une différence en faveur du placebo avec moins d'effets indésirables

généraux et spécifiques en comparaison avec l'amitriptyline. A priori, ces effets indésirables ne sont pas responsables de plus d'abandons dans le groupe amitriptyline que dans le groupe placebo.

Selon la méthode REB aucun de nos essais n'est considéré comme concluant car aucune de nos études ne remplit les deux critères suivants :

- Avoir un risque de biais global faible
- Avoir un résultat statistiquement significatif pour les critères de jugements définis à priori (ou ayant été modifiés au cours de l'essai sans avoir connaissance des résultats) et après avoir pris en compte le risque d'inflation du risque alpha en cas de multiplicité de tests et l'erreur de type I.

Par ailleurs, nous n'avons pas retrouvé d'articles supplémentaires non publiés sur la plateforme ClinicalTrials.gov traitant du même sujet, ce qui pourrait suggérer l'absence de biais de publication.

Selon la méthode REB il existe donc un « manque de preuve » concernant l'efficacité de l'amitriptyline dans les céphalées de tension chroniques. Selon la méthode GRADE le niveau de preuve est considéré comme étant « très faible ».

2. Confrontation à la littérature :

Nous avons retrouvé une revue systématique s'intéressant à l'utilisation de l'amitriptyline dans les céphalées de tension chroniques. Celle-ci a été réalisée en 2008 par [Eulalia Torrente Castells](#), [Eduardo Vázquez Delgado](#) et [Cosme Gay Escoda](#).

Dans cette revue de la littérature, 10 études ont été incluses, parmi celles-ci six ne font pas partie de notre revue de la littérature, il s'agit de :

- [Nappi et al. \(1990\) \(22\)](#): A new 5-HT₂ antagonist (ritanserin) in the treatment of chronic headache with depression. A double-blind study vs amitriptyline.
- [Mitsikostas et al. \(1997\) \(23\)](#): Buspiron vs amitriptyline in the treatment of chronic tension-type headache.
- [Holroyd et al. \(2001\) \(24\)](#): Management of chronic tension-type headache with tricyclic antidepressant medication, stress management therapy, and their combination: a randomized controlled trial.

- [Rampello et al. \(2004\) \(25\)](#): Evaluation of the prophylactic efficacy of amitriptyline and citalopram, alone or in combination, in patients with comorbidity of depression, migraine, and tension-type headache.
- [De Tommaso et al. \(2005\) \(26\)](#): Effects of amitriptyline and intra-oral device appliance on clinical and laser-evoked potentials features in chronic tension-type headache.
- [Bettucci et al. \(2006\) \(27\)](#): Combination of tizanidine and amitriptyline in the prophylaxis of chronic tension-type headache: evaluation of efficacy and impact on quality of life.

Nous avons exclu les études de Mitsikostas et al. (1997), Nappi et al. (1990), Bettucci et al. (2006), De Tommaso et al. (2005), Rampello et al. (2004) car elles n'évaluent pas l'amitriptyline contre un placebo.

Dans l'étude de Holroyd et al. (2001), les participants ont été randomisés pour recevoir soit :

- un antidépresseur tricyclique (chlorhydrate d'amitriptyline, jusqu'à 100 mg/j, ou chlorhydrate de nortriptyline, jusqu'à 75 mg/j)
- un placebo,
- une thérapie de gestion du stress (par exemple, relaxation, adaptation cognitive) (3 séances et 2 contacts téléphoniques) et un placebo,
- une thérapie de gestion du stress et un médicament antidépresseur.

Nous avons choisi d'exclure cette étude car dans cette étude le groupe "amitriptyline" se compose de patients ayant été traités par amitriptyline mais également par des patients ayant mal toléré l'amitriptyline et donc traités avec de la nortriptyline par la suite. Les résultats sont mesurés pour tout le groupe sans faire de distinction entre les deux molécules. De ce fait cette étude n'évalue pas l'efficacité de l'amitriptyline dans les céphalées de tension mais elle évalue l'efficacité des antidépresseurs tricycliques dans les céphalées de tension.

L'analyse de ces 10 essais a permis aux auteurs de conclure que "bien que les résultats obtenus n'aient pas toujours été statistiquement significatif pour tous les paramètres cliniques, l'amitriptyline (10 à 100 mg/jour) peut être considéré comme ayant démontré l'efficacité dans le traitement prophylactique de la céphalée de tension chronique." Les auteurs ont évalué le niveau de preuve selon les critères SORT (Strength of Recommendation Taxonomy) : 5 articles montraient un niveau de preuve scientifique classé 2 et 5 autres articles un niveau de preuve scientifique classé 1. Les auteurs concluent finalement à une recommandation de type A en faveur de l'utilisation de l'amitriptyline dans le traitement des céphalées de tension chroniques.

Notre étude se différencie de cette étude dans le sens où notre méthodologie intègre les risques de biais dans la détermination du niveau de preuve et de ce fait nous concluons à un niveau de preuve “très faible”.

3. Limites :

Nous avons rencontré plusieurs limites dans la réalisation de cette revue de la littérature :

- Peu d'études ont été incluses dans notre revue de la littérature. Cela s'explique par le fait que nous comparons l'amitriptyline à un placebo, hors de nombreuses études comparent l'amitriptyline uniquement à une molécule active. Le plus souvent dans ces études l'amitriptyline est considérée comme le gold standard.
- Aucune méta-analyse n'a pu être réalisée du fait de l'hétérogénéité dans la mesure des résultats mais également du manque de données rapportées par les auteurs. Nous avons essayé de joindre les différents auteurs pour récupérer les données brutes mais aucun auteur n'a répondu à notre demande.
- Les auteurs ont utilisé des unités de mesures et des échelles différentes pour leur critère de jugement, mais également des scores qui étaient non reproductibles entre les études. Ces différences rendent l'interprétation des résultats un peu plus approximative.
- Plusieurs études évaluent l'efficacité d'une autre molécule en priorité en même temps que l'amitriptyline. Dans la mesure où cela n'est pas l'objectif principal de l'étude, cela peut avoir un impact sur l'interprétation du résultat.
- Plusieurs de nos études sont très anciennes, la plus récente datant de 2004. La rédaction des articles peut être parfois succincte avec peu de détails méthodologiques rapportés.
- Nos études sont toutes à petits effectifs et donc de faible puissance, le plus grand effectif étant celui de l'étude de Pfaffenrath 1994 avec au total 197 patients.
- Dans l'évaluation de la douleur l'effet placebo peut être important, cela peut avoir un effet confondant sur les résultats d'autant plus quand les études sont de faible puissance et donc biaiser les conclusions de notre étude.
- Nous avons choisi de sélectionner uniquement les articles en français ou en anglais. De ce fait un article dans une autre langue a été exclus. Cela peut potentiellement biaiser les résultats de notre étude.
- Nos conclusions ne peuvent pas s'appliquer à une population pédiatrique car nos études incluent uniquement des adultes.

- Bien que nous n'ayons pas retrouvé d'article supplémentaire complet sur la plateforme ClinicalTrials.gov, un biais de publication ne peut être exclu de manière empirique.
- Trois études ont été soutenues financièrement par l'industrie pharmaceutique (fondation Lundbeck et fondation pharmaceutique 1991).

4. Implications pour la pratique :

Il apparaît dans notre étude que la preuve de l'efficacité de l'amitriptyline est "très faible" selon GRADE et qu'elle "manque de preuve" selon REB. Nos résultats sont donc en défaveur de l'utilisation de l'amitriptyline dans cette indication bien qu'il semble que l'amitriptyline ait un effet positif dans le traitement des céphalées de tension chroniques. L'amitriptyline est recommandée comme traitement de première intention par l'ANAES en France (4) (Agence Nationale d'Accréditation et d'Evaluation en Santé), ainsi que par l'EFNS (European Federation of the Neurological Societies) (5) et la société suisse des céphalées (6). L'amitriptyline est reconnue par de nombreux cliniciens et auteurs comme le gold standard du traitement des céphalées de tension, or notre étude démontre qu'il n'existe pas de preuve scientifique solide attestant de son efficacité. Elle ne devrait donc plus être considérée comme le gold standard par la communauté scientifique et sa place dans le traitement des céphalées devrait être remise en cause.

Pour la pratique, il convient donc d'utiliser cette molécule avec précaution chez les patients en mettant en balance les bénéfices et les risques et de fournir au patient une information claire concernant l'efficacité et les effets indésirables.

5. Implications pour la recherche :

La prise en compte des biais étant centrale dans la méthode REB et ayant également une place importante dans l'appréciation du niveau de preuve GRADE, nos résultats sont en défaveur de l'utilisation de l'amitriptyline dans cette indication bien qu'il semble que l'amitriptyline ait un effet positif dans le traitement des céphalées de tension chroniques.

L'amitriptyline est une molécule pourvoyeuse d'effets secondaires spécifiques facilement identifiables, ce qui met en péril le double aveugle dans la plupart des études. Nous préconisons donc la réalisation de nouvelles études évaluant l'amitriptyline contre un placebo et contre une molécule ayant des effets indésirables proches de l'amitriptyline afin de maintenir le double aveugle. La mirtazapine semble être la molécule de choix. En effets les études montrent que l'amitriptyline et la mirtazapine sont toutes les deux aussi efficaces pour réduire la fréquence et l'intensité des céphalées de tension chroniques versus placebo (27). Par ailleurs la mirtazapine présente également des effets indésirables de type somnolence et sècheresse buccale mais dans une moindre mesure.

Nous constatons également un pourcentage élevé d'abandon dans les études sans que cela ne soit en lien avec les effets indésirables de l'amitriptyline. De plus l'effet placebo étant important dans l'évaluation de la douleur, il est important de réaliser des études avec des effectifs importants afin d'avoir une puissance suffisante pour mettre en évidence une différence significative. Afin de minimiser les risques liés aux données manquantes nous préconisons l'utilisation de méthodes de gestion des données manquantes efficaces.

Seule l'étude de Pfaffenrath suit les patients sur une durée de 6 mois. Les autres études n'évaluent pas l'efficacité de l'amitriptyline sur le long terme, or les personnes souffrant de céphalées chroniques sont amenées à prendre le médicament sur de longues périodes. Il serait intéressant d'évaluer l'efficacité de l'amitriptyline sur les moyens et longs termes.

VI- CONCLUSION

Les céphalées de tension chronique sont difficiles à traiter car ce sont encore des entités mal connues. En effet leur physiopathologie n'est pas complètement élucidée. L'hypothèse la plus probable est une contraction myofaciale anormale de certains muscles induisant une hypoxie musculaire. S'ensuit alors une libération de médiateurs de la douleur. L'exposition répétée et prolongée à ces neuromédiateurs modifierait la sensibilité à la douleur au niveau du système nerveux périphérique et central contribuant à la chronicisation des douleurs. Quelques études explorent l'hypothèse d'une participation de facteurs génétiques, mais actuellement aucun gène n'a été identifié.

De par ses multiples mécanismes d'action, l'amitriptyline possède de nombreuses propriétés : antidépressive, analgésique, sédatrice et anticholinergique. Elle semble ainsi être une molécule de choix pour le traitement des céphalées de tension chroniques. Elle est d'ailleurs recommandée comme traitement de première intention pour cette indication dans plusieurs pays dont la France.

Notre recherche bibliographique ne nous a pas permis d'identifier de méta analyse traitant du sujet. Cependant nous avons trouvé une revue systématique de la littérature datant de 2008 qui conclue à un niveau de preuve SORT de niveau A. Cette revue de la littérature incluait en majorité des études qui ne comparaient pas l'amitriptyline à un placebo. Nous avons donc souhaité réaliser une revue systématique de la littérature selon la méthode REB (Rebuild the Evidence Base) pour évaluer l'efficacité de l'amitriptyline en ne sélectionnant uniquement que des essais contrôlés contre placebo, randomisés et en double aveugle. La méthodologie REB vise à vérifier le niveau de preuve de tous les médicaments utilisés en médecine courante et pour chacune de leurs indications, selon une méthodologie stricte. Ce travail a pour but d'aider les médecins dans la prescription des thérapeutiques en apportant une information claire concernant l'efficacité du médicament dans une indication donnée.

Cette revue systématique inclut sept études. Les patients ont été suivis entre 4 semaines et 6 mois, au total 792 patients ont été évalués. Toutes nos études sont à risque élevé de biais selon l'outil RoB 2. Cinq études présentent des critères de jugements principaux multiples et parmi celles-ci 4 ont utilisé une méthode d'ajustement du risque alpha. Nous n'avons pas pu réaliser de méta-analyse dans la mesure où il existe une hétérogénéité des unités de mesure, des scores ou des échelles utilisées par les différents auteurs. Dans certaines études, le détail des effectifs

et les moyennes n'étaient pas présentés. Nous avons essayé de contacter les auteurs pour récupérer les données brutes. Certains auteurs nous ont répondu mais malheureusement à chaque fois les données n'étaient plus disponibles. Les résultats de notre étude suggèrent une différence en faveur de l'amitriptyline sur la durée, l'intensité et la fréquence des céphalées, avec des résultats plus ou moins significatifs selon les études. Selon la méthode REB nous avons conclu qu'il existait "un manque de preuve" et selon GRADE le niveau de preuve est "très faible ". Nous considérons que l'utilisation de l'amitriptyline actuellement en tant que gold standard n'est pas légitime du fait d'un manque de preuves scientifiques solides concernant son efficacité pour cette indication.

La méthode REB est très sensible aux risques de biais, il en est de même pour la méthode GRADE, ce qui explique ces résultats. Il y a un nombre faible d'études dans la littérature, toutes sont anciennes, avec de petits effectifs et elles détaillent très peu leur méthodologie. Par ailleurs, l'amitriptyline est une molécule pourvoyeuse d'effets indésirables spécifiques tels que la sécheresse buccale et la somnolence, rendant le maintien du double aveugle difficile lors des essais. Ces effets indésirables ne sont a priori pas responsables d'abandons plus importants, mais nous remarquons qu'il existe dans les études de forts taux de données manquantes. Il serait donc intéressant de réaliser de nouvelles études de forte puissance, contre placebo mais également contre une molécule pourvoyeuse des mêmes effets indésirables spécifiques que ceux de l'amitriptyline, sur de plus longues périodes, en utilisant une méthode efficace de gestion des données manquantes.

VII- FINANCEMENT ET DÉCLARATION DE CONFLITS D'INTÉRÊTS

Les auteures déclarent ne pas avoir reçu de financement pour la réalisation de cette revue systématique de la littérature et ne pas avoir de conflit d'intérêt.

VIII- BIBLIOGRAPHIE

1. The International Classification of Headache Disorders, 3rd edition (beta version). Cephalalgia. 2013;33(9):629-808.
2. Linde M, Gustavsson A, Stovner LJ, Steiner TJ, Barré J, Katsarava Z, et al. The cost of headache disorders in Europe: the Eurolight project. European Journal of Neurology. 201, 19 (5:703-711.
3. Panzeri M, Ryvlin P, Staeger P, Gautschi R, Amstutz V. Approche myofasciale dans la prise en charge des céphalées de tension : ce que dit la science. Revue Médicale Suisse. 2020;16(687) : 600-605
4. Gantenbein, Andreas & Palla, Antonella & Sturzenegger, Mathias. (2020). Recommandations thérapeutiques de la SSC pour les céphalées et •algies faciales. Forum Médical Suisse – Swiss Medical Forum. 10.4414/fms.2020.08466.
5. Bendtsen, L., Evers, S., Linde, M., Mitsikostas, D.D., Sandrini, G. and Schoenen, J. (2010), EFNS guideline on the treatment of tension-type headache – Report of an EFNS task force. European Journal of Neurology, 17: 1318-1325.
6. Haute Autorité de Santé HAS. CCQ (céphalées chroniques quotidiennes) : diagnostic, rôle de l'abus médicamenteux, prise en charge [Internet]. Saint-Denis La Plaine; 2004. Disponible sur: https://www.has-sante.fr/jcms/c_272365/fr/ccq-cephalees-chroniques-quotidiennes-diagnostic-role-de-l-abus-medicamenteux-prise-en-charge.
7. Leucht C, Huhn M, Leucht S. Amitriptyline versus placebo for major depressive disorder. Cochrane Database of Systematic Reviews 2012, Issue 12. Art. No.: CD009138.
8. Amitriptyline : substance active à effet thérapeutique [Internet]. VIDAL. Disponible sur: <https://www.vidal.fr/medicaments/substances/amitriptyline-6678.html>.
9. LAROXYL 25 mg, comprimé pelliculé. Résumé des caractéristiques du produit. Dans : Base de données publique des médicaments [Internet]. [cité 11 févr 2022]. Disponible sur : <https://base-donnees-publique.medicaments.gouv.fr/affichageDoc.php?specid=65892928&typedoc=R>.
10. Torrente Castells, E., Vázquez Delgado, E., & Gay Escoda, C. (2008). Use of amitriptyline for the treatment of chronic tension-type headache. Review of the literature. Medicina oral, patologia oral y cirugía bucal, 13(9), E567–E572.
11. Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews BMJ 2021; 372 :n71 doi:10.1136/bmj.n71

12. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; 366: l4898.
13. Boussageon R, Blanchard C, Menini T, Pereira B, Naudet F, Kassai B, et al. Rebuild the Evidence Base (REB) project: a method for reviews and meta-analyses focusing on benefit-risk assessment of interest to patients and general practitioners.
14. Schünemann H, Brožek J, Guyatt G, Oxman. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. handbook [Internet]. Disponible sur: <https://gdt.grade.pro.org/app/handbook>
15. /handbook.html.

Articles inclus dans notre étude

16. Lance JW, Curran DA. Treatment of chronic tension headache. *The Lancet*. 1964;283(7345):1236-1239.
17. Göbel H, Hamouz V, Hansen C, Heininger K, Hirsch S, Lindner V, Heuss D, Soyka D. Chronic tension-type headache: amitriptyline reduces clinical headache-duration and experimental pain sensitivity but does not alter pericranial muscle activity. *Pain*. 1994;59(2):241-249.
18. Pfaffenrath V, Diener HC, Isler H, Meyer C, Scholz E, Taneri Z, Wessely P, Zaiser-Kaschel H, Haase W, Fischer W. Efficacy and tolerability of amitriptyline in the treatment of chronic tension-type headache: a multi-centre controlled study. *Cephalalgia*. 1994;14(2):149-155.
19. Bendtsen 1996 L, Jensen R, Olesen J. A non-selective (amitriptyline), but not a selective (citalopram), serotonin reuptake inhibitor is effective in the prophylactic treatment of chronic tension-type headache. *J. Neurol. Neurosurg. Psychiatry*. 1996;61(3):285-290.
20. Diamond S, Baltes BJ. Chronic tension headache--treated with amitriptyline--a double-blind study. *Headache*. 1971;11(3):110-116.
21. Bendtsen L, Jensen R. Amitriptyline reduces myofascial tenderness in patients with chronic tension-type headache *Cephalalgia*. 2000;20(6):603-610.
22. Ashina S, Bendtsen L, Jensen R. Analgesic effect of amitriptyline in chronic tension-type headache is not directly related to serotonin reuptake inhibition. *Pain*. 2004;108(1-2):108-114.

Articles exclus de notre étude

23. Nappi G, Sandrini G, Granella F, Ruiz L, Cerutti G, Facchinetti F, et al. A new 5-HT₂ antagonist (ritanserin) in the treatment of chronic headache with depression. A double-blind study vs amitriptyline. *Headache*. 1990 Jun;30(7):439-44.
24. Mitsikostas DD, Gatzonis S, Thomas A, Ilias A. Buspirone vs amitriptyline in the treatment of chronic tension-type headache. *Acta Neurol Scand*. 1997 Oct;96(4):247-51.
25. Holroyd KA, O'Donnell FJ, Stensland M, Lipchik GL, Cordingley GE, Carlson BW. Management of chronic tension-type headache with tricyclic antidepressant medication, stress management therapy, and their combination: a randomized controlled trial. *JAMA*. 2001 May 2;285(17):2208-15.
26. Rampello L, Alvano A, Chiechio S, Malaguarnera M, Raffaele R, Vecchio I, et al. Evaluation of the prophylactic efficacy of amitriptyline and citalopram, alone or in combination, in patients with comorbidity of depression, migraine, and tension-type headache. *Neuropsychobiology*. 2004;50(4):322-8.
27. De Tommaso M, Shevel E, Libro G, Guido M, Di Venere D, Genco S, et al. Effects of amitriptyline and intra-oral device appliance on clinical and laser-evoked potentials features in chronic tension-type headache. *Neurol Sci*. 2005 May;26 Suppl 2:s152-4.
28. Bettucci D, Testa L, Calzoni S, Mantegazza P, Viana M, Monaco F. Combination of tizanidine and amitriptyline in the prophylaxis of chronic tension-type headache: evaluation of efficacy and impact on quality of life. *J Headache Pain*. 2006 Feb;7(1):34-6.
29. Ghadiri-Sani M, Silver N. Headache (chronic tension-type). *BMJ Clin Evid*. 2016;2016:1205.

IX- ANNEXES

1. J.W Lance 1964: Treatment of chronic tension headache

a) Caractéristiques des études

Méthode	2 essais : les deux sont prospectifs, randomisés, en double aveugle, contrôlés par un placebo, en crossover
Participants	1ère partie : N = 280 ; 2ème partie : N = 27 Adultes suivis à la clinique de céphalées Hôpital prince Henry, présentant des céphalées de tension depuis plus d'un an, récurrentes avec, constantes, pressantes ou serrées et n'étant habituellement pas associées à des caractéristiques de migraine
Interventions	<u>1 ère partie</u> : Les participants ont reçu soit des sédatifs (amylobarbitone), soit des Vasodilatateurs, soit la combinaison d'un sédatif et d'un médicament vasodilatateur, soit du "Ballergal " un mélange de phénobarbital, ergotamine et belladone, soit du Methysergide un dérivé d'acide lysergique, soit du citrate d'orphenadrine, soit du chlordiazepoxide, soit du Diazépam, soit de l'Amitriptyline soit de l'Imipramine, soit du "Parstelin" un mélange de tranlycypromine et trifluopérazine, soit un Placebo Si les patients signalaient une amélioration significative des céphalées lors de la visite mensuelle, ils poursuivaient le même traitement. Si les patients signalaient l'absence d'amélioration ou une légère amélioration lors de la visite mensuelle, les patients étaient dirigés vers un nouveau traitement et la procédure se répétait. <u>2 ème partie</u> : Les participants ont reçu consécutivement soit un traitement placebo puis de l'amitriptyline soit vice versa un traitement par amitriptyline puis un placebo.
Outcomes	<u>Le critère de jugement principal était dans la première partie</u> : " pratiquement la disparition des céphalées", "plus de 50% d'amélioration" ou " pas de modification ". <u>Dans la seconde partie</u> : "amélioration" ou "absence d'amélioration".

Notes	Exclusions : Etaient exclus de l'étude les patients présentant des céphalées occasionnelles ou peu fréquentes. Etude monocentrique Durée de l'essai : 1ère partie (N = 280) :12 mois 2ème partie (N = 27) : 2 mois Analyse en ITT Absence d'information concernant les perdus de vue <u>Première partie :</u> Résultat significatif pour le groupe Amitriptyline versus placebo <u>Seconde partie :</u> Résultat significatif pour l'Amitriptyline (15 améliorations) versus Placebo (3 améliorations)
-------	--

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« They were then given one of a group of drugs and instructed to report back to the clinic in a month. »	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	« Neither patient nor physician knew the nature of the tablets, the key being retained by hospital pharmacy until the end of the trial »	PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« Patients in each group initially had headache comparable duration and frequency. » Pas d'autre information sur les caractéristiques de base des groupes.	NI
Evaluation du risque de biais		Quelques inquiétudes

Domaine lié au cross-over : Biais dû à l'effet « carry-over »		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	Initialement oui, au début 12 groupes de 27 participants. Ensuite pas d'information sur l'évolution des groupes (perdus de vue, maintien dans un groupe...) mais déséquilibre certain	PN
S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?	Pas de période de wash-out, pas d'analyse des résultats en fonction du temps	NI
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	Absence d'effet carry-over car mesures réalisées 4 semaines après le changement de molécule et : <u>Pour l'amitriptyline</u>, : - du fait d'une demi-vie de entre 22 et 40 heures (source Résumé des caractéristiques du produit - ELAVIL 25 mg, comprimé pelliculé - Base de données publique des médicaments (medicaments.gouv.fr)) et du constat par les auteurs de l'article d'une récurrence des symptômes de céphalées dans les 2 à 7 jours suivant son arrêt « the headache returned two to seven days after [...] amitriptyline was stopped », on peut raisonnablement penser qu'au bout d'une semaine il n'y a plus d'effet carry-over <u>Pour l'imipramine</u> : demi-vie entre 9 et 20h (source : http://agence-prd.ansm.sante.fr/), sans effet persistant connu sur le corps après l'élimination de la molécule. <u>Pour le diazepam</u> ½ vie 32 à 47 heures, pas d'effet une fois la molécule définitivement éliminée du corps <u>Pour le Clordiazepoxide</u> ½ vie : 6 à 30 heures, mais métabolites actifs (oxazepam) qui allongent la demi vie. Pas d'effet une fois la molécule et ses métabolites éliminés (http://agence-prd.ansm.sante.fr/) Pour le Parstelin (tranylcypromine et trifluopérazine) : l'association ne semble plus être commercialisée, et n'a pas apporté de preuve de son efficacité sur les céphalées de tension, donc pas d'effet carry-over attendu. <u>Pour le béllergal</u> (phénobarbitone + ergotamine tartrate + belladonna alkaloids, à priori n'est plus commercialisé (donc difficulté d'accès aux informations), mais d'une part la demi-vie du médicament semble être liée au phénobarbital qui	PY

	a une demi-vie de 53 à 118 heures) n'a pas montré son efficacité dans les céphalées de tension. Donc pas d'effet carry-over attendu <u>Pour l'orphénadrine</u> : à priori n'a pas fait ses preuves dans le traitement des céphalées de tension, donc pas d'effet rémanent prévisible <u>Pour le méthysergide</u> : demi-vie 4 heures, pas d'effet après son élimination totale du corps. <u>Pour les vasodilatateurs</u> : l'acide nicotinique : ce médicament n'est plus commercialisé, les données n'ont pu être trouvées et pour l'hydrogenated ergot alkaloid « hydergine » la demi-vie est de 13 à 15 heures. Le médicament n'a à notre connaissance à ce jour pas montré d'efficacité dans la prise en charge des céphalées de tension. Donc dans notre étude, son effet carry-over peut être considéré comme nul. <u>Pour les sédatifs</u> : amylobarbitone : d'après nos recherches le médicament n'est actuellement plus commercialisé et nous n'avons pu trouver sa demi-vie. La molécule la plus proche actuellement sur le marché est le phénobarbital, $\frac{1}{2}$ vie : 50 à 140 heures (Source : base-donnees-publique.medicaments.gouv.fr). Le médicament n'est pas connu pour avoir un effet persistant après son élimination du corps.	
Evaluation du risque		Quelques inquiétudes

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Pas d'informations sur les effets indésirables mais plusieurs molécules testées avec de nombreux effets secondaires plus ou moins spécifiques. « Neither patient nor physician knew the nature of the tablets, the key being retained by hospital pharmacy until the end of the trial »	PN
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PN
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		Non concerné

2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITT	
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		
Evaluation du risque de biais		Risque faible

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	Pas d'information	NI
3.2 Existe-t-il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Essai de minimisation par la règle « Any patient who failed to report for follow-up was classified as unchanged on the last treatment prescribed. »	PY
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?		
3.4 Si Y / PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?		
Evaluation du risque de biais		Risque faible

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un		PN

groupe d'intervention à l'autre ?		
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Evaluateur = patient Double aveugle a priori maintenu	PN
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?		Non concerné
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?		Non concerné
Evaluation du risque de biais		Risque faible

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information sur un enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
5.4 : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global selon RoB :

Biais	Jugement par domaine de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Séquence de randomisation non décrite Peu d'informations concernant les caractéristiques de base des participants
Biais dû à l'effet « carry-over »	Quelques inquiétudes	Nombre de participants dans chaque bras à chaque période déséquilibré Absence de période de wash-out Absence d'effet carry-over identifié cependant
Biais dû aux données manquantes	Risque faible	Double aveugle maintenu Analyse en ITT
Biais dû aux données manquantes	Risque faible	Absence d'informations concernant le pourcentage de données manquantes Minimisation du risque de biais par la règle : chaque donnée manquante pour une molécule remplacée par "absence d'amélioration"
Biais dû à des erreurs de mesure dans les résultats	Risque faible	Méthode de mesure appropriée Double aveugle maintenu
Biais dans la sélection des résultats	Quelques inquiétudes	Absence de temporalité du protocole Absence de sélection des résultats Une seule analyse réalisée Résultats de toutes les périodes disponibles
Autres biais		Possible biais de confusion : absence d'exclusion entre autres des patients ayant déjà pris les molécules testées, absence d'exclusion des patients sous autres médications psychotropes, des patients présentant des addictions, des pathologies psychiatriques lourdes... Absence d'informations concernant les conflits d'intérêts Pas de méthode d'ajustement rapportée car un seul critère de jugement
Risque global de biais	Risque élevé	

d) Tableau de résultats :

Auteur : Lance Année : 1964						
1ère partie: Essai randomisé en cross over Nombre de participants : N = 80 Analyse en ITT 12 bras d'intervention :						
sédatifs (amylobarbitone),	Vasodilatateurs	combinaison d'un sédatif et d'un médicament vasodilatateur	"Ballergal " un mélange de phénobarbital, ergotamine et belladone	Methysergide un dérivé d'acide lysergique	citrate d'orphénadrine	
chlordiazepoxide	Amitriptyline	Diazépam	Imipramine	"Parstelin" un mélange de tranylcypromine et trifluopérazine	Placebo	
2ème partie : Essai randomisé en crossover 2 bras d'intervention : Amitriptyline et Placebo Nombre de participants N = 27 Analyse en ITT						
	Critère de jugement	Mesure du critère	Valeur rapportée	Amitriptyline	Placebo	Significatif ou non versus placebo
Critère de jugement primaire :	2ème essai N = 27 : Amélioration des symptômes (Fréquence ou intensité) A la fin du premier mois	Disparition des céphalées	Nombre de patients	9	1	Significatif par méthode « fourfold contingency table (Fisher 1961) et « exact
		Amélioration supérieur de 50% céphalées		15	5	
		Inchangée		11	12	
		Perdu de vue		1	2	

		Amélioration totale et > 50%	Pourcentage	67	30	« méthode » par distribution binomiale
	Premier Essai N = 280 : Amélioration des symptômes (fréquence ou intensité) à la fin de l'essai	Disparition des céphalées	Nombre de patients (Au total 104 patients traités par AT au cours de l'essai et 49 pour placebo)	22	2	
		Amélioration supérieur de 50% céphalées		36	9	
		Inchangée		40	34	
		Perdu de vue		7	4	
		Amélioration totale et > 50%	Pourcentage	55	22	

2. Gobel 1994: Chronic tension-type headache: amitriptyline reduces clinical headache duration and experimental pain sensitivity but does not alter pericranial muscle activity readings

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles
Participants	N =78, adultes de 18 à 75 ans, non hospitalisés, présentant des céphalées de tension au moins 15 jours par mois durant les 6 derniers mois, sans traitement prophylactique pour ces céphalées durant les 6 mois précédant l'essai.
Interventions	Les participants ont reçu soit : - 25 mg d'hydrochlorure d'amitriptyline en libération prolongé : le dosage était progressivement augmenté : 1 gélule la première semaine, 2 gélules (50mg) la seconde et 3 gélules (75mg) à partir de la 3ème semaine. Soit - Un placebo dans une gélule identique, le soir avant le coucher.
Outcomes	<u>Critère de jugement principal :</u> La durée journalière moyenne des céphalées par semaine. <u>D'autres paramètres neurophysiologiques ont été enregistrés :</u> Activité EMG, la suppression extéroceptive de l'activité du muscle temporal, la variation négative du contingent (CNV) et l'expérience de la sensibilité à la douleur Les effets secondaires La consommation d'analgésiques
Notes	Critères d'exclusion : patientes enceintes ou ayant un projet de grossesse, les situations contre indiquant la prise d'amitriptyline (ex épilepsie, glaucome, HTA, dysfonction prostatique, troubles cardiaques ou hépatiques), les patients présentant une addiction médicamenteuse/toxicomanie, les antécédents de maladies psychiatriques définies dans le DSM III, les patients présentant d'autres formes de céphalées primaires ou symptomatiques, la prise de médicaments psychoactifs dans les 6 mois précédant l'essai et pendant l'essai, la non-participation au rendez-vous de suivi, les patients ayant des antécédents de réaction allergique ou d'effets indésirables au médicament, l'abus de prise médicamenteuse pour traiter la céphalée de tension c'est à dire plus de 30 grammes d'aspirine par mois ou équivalent d'autres analgésiques. Durée de l'essai : 6 semaines Analyse en per protocole

	Participants analysés N = 53 Perdus de vue / exclus : 25 patients dont 4 pdv, 21 exclus pour violation du protocole Résultat significatif pour le groupe amitriptyline versus placebo à la sixième semaine de traitement seulement.
--	--

b) Tableau RoB 2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	<i>"The study was designed as a double blind randomized and placebo controlled parallel group"</i>	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	"The patients were assigned in a double-blind and randomized fashion to either placebo or the amitriptyline group"	NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	<i>"The baseline demographic and clinical headache data are summarized in table 1. These data revealed no significant differences between the amitriptyline and placebo groups. Some 86 % of the placebo group and 71% of the amitriptyline group had suffered from chronic tension -type headache for more than 5 years (X2 test : not significant)."</i> Cependant données concernant la population finale (N = 53) et non la population initiale.	NI
Evaluation du risque de biais		Quelques inquiétudes

Domaine 2 : Biais dû aux écarts par rapport aux interventions envisagées (effet de l'adhésion à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	The study medication was 25 mg amitriptyline HCl in a sustained release capsule [...] or placebo in an identical capsule “The side-effect rates within each group and between the placebo group and the amitriptyline group did not differ significantly during the first to the 6th week treatment. “ Cependant les effets secondaires spécifiques de l'amitriptyline deux fois plus marqués de façon significative dans le groupe traité que dans le groupe control : Dryness of the mouth (amitriptyline 54% placebo 17% X ² test p < 0.05 Mild drowsiness amitriptyline 62% placebo 27%, X ² test p < 0.05	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Des interventions non protocolaires importantes étaient-elles équilibrées entre les groupes d'intervention ?		Non concerné
2.4 [Le cas échéant] Y a-t-il eu des échecs dans la mise en œuvre de l'intervention qui auraient pu affecter les résultats ?		PN
2.5 [Le cas échéant] Y a-t-il eu un non-respect du schéma d'intervention assigné qui aurait pu affecter les résultats des participants	21 exclus pour violation du protocole sur 78 patients => 26.9% population initiale de l'étude Failure to comply with regular intake of medication No attendance of follow up examination Failure to complete the headache diary	PY
2.6 Si N / PN / NI à 2.3 ou Y /PY/NI à 2.4 ou 2.5 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet du respect de l'intervention	Analyse en per protocole	PN
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	25 patients dont 4 pdv pour effets indésirables et 21 exclus pour violation du protocole. 25 sur 78 => 32% de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?		PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	26.9% de données manquantes, en lien avec le résultat final : Effets indésirables et manque de compliance	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	The frequency of causes of drop out did not differ significantly between the 2 groups (X ² test : p>0.05) Mais drop out en lien avec l'étude : effets indésirables, manque de compliance... et concernant plus d'un quart de la population initiale de l'étude.	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Double aveugle non maintenu Evaluateurs : médecin et/ou patient	PY
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critères de jugements subjectifs	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Absence d'éléments permettant de statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas de base de données d'enregistrement des protocoles lors de la réalisation de l'étude	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Pas d'information concernant la technique employée pour la randomisation. Données démographiques de la population finale présentée et non de la population initiale.
Biais dû aux écarts par rapport aux interventions envisagées (effet de l'adhésion à l'intervention)	Risque élevé	Double aveugle a priori non maintenu (Effets indésirables spécifiques de l'amitriptyline) Manque d'adhésion à l'intervention Pas d'analyse appropriée pour estimer l'effet de l'adhésion.
Biais dû aux données manquantes	Risque élevé	25 patients n'ont pas été inclus dans l'évaluation car l'ensemble des données étaient incomplètes : 4 ont présentés des effets secondaires et 21 n'ont pas respectés le protocole de l'étude. Potentiel impact sur la vraie valeur bien que la fréquence des causes d'abandon ne différait pas significativement entre les deux groupes.
Biais dû à des erreurs de mesure dans les résultats	Risque élevé	Méthode de mesure adaptée Double aveugle à priori non maintenu Critères de jugements subjectifs
Biais dans la sélection des résultats	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement.
Autres biais		Pas d'informations sur les conflits d'intérêts. Etude monocentrique Pas de méthode d'ajustement du risque alpha rapportée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Gobel année : 1994					
Etude randomisée parallèle en 2 bras : Amitriptyline vs Placebo N = 78, adultes de 18 à 75 ans Patients randomisés : Amitriptyline (A) n = 24, Placebo (P) n = 29 Durée : 6 semaines Analyse en ITTm					
	Critère de jugement	Mesure du critère et de la valeur rapportée (moyenne $\pm$ écart type)	Amitriptyline		Significatif ou non contre placebo
Critère de jugement principal	Durée de la céphalée	Moyenne d'heures de céphalées quotidienne par semaine	Durée journalière en heure / Semaine	Amélioration intragroupe /S	
			Semaine 1 P : 9.9 $\pm$ 5.1 A : 11.1 $\pm$ 5.0		Non significatif
			Semaine 2 P : 9.7 $\pm$ 5.6 A : 10.7 $\pm$ 4.7	Semaine 2 - 0.19 $\pm$ 3.6 - 0.43 $\pm$ 1.9	p = 0.820 NS
			Semaine 3 P : 8.7 $\pm$ 5.3 A : 9.5 $\pm$ 5.6	Semaine 3 - 1.1 $\pm$ 3.4 - 1.6 $\pm$ 3.1	p = 0.559 NS
			Semaine 4 P : 9.2 $\pm$ 5.6 A : 9.0 $\pm$ 5.5	Semaine 4 - 0.72 $\pm$ 3.5 - 2.2 $\pm$ 3.6	p = 0.147 NS
			Semaine 5 P = 9.3 $\pm$ 5.3 A = 8.5 $\pm$ 6.1	Semaine 5 - 0.59 $\pm$ 3.3 - 2.6 $\pm$ 4.2	p = 0.068 NS
			Semaine 6 P : 9,6 $\pm$ 6,0 A : 7,9 $\pm$ 6.0	Semaine 6 - 0.28 $\pm$ 3.0 - 3.2 $\pm$ 4.1	p = 0.007
			Amélioration intragroupe cumulée entre S1 et S4-6 Placebo 1.6 $\pm$ 8.67h Amitriptyline 7.82 $\pm$ 11h		p < 0.01 Mann-Whitney U test

Critères de jugement secondaires			Pré-traitement	Post-traitement	
	La suppression extéroceptive de l'activité du muscle temporal par EMG	EMG durant relaxation	A : 11.2 ± 15.6 P : 10.5 ± 10.6	A : 10.8 ± 12.9 P : 10.7 ± 9.4	Non significatif
		EMG durant effort maximal	A : 185.4 ± 127.9 P : 157.2 ± 110.0	A : 176.2 ± 122.4 P : 169.9 ± 133.5	Non significatif
		Durée de ES2	A : 8.5 ± 12.7 P : 17.2 ± 21.3	A : 9.5 ± 5.4 P : 16.3 ± 20.1	Non significatif
	La variation négative du contingent (CNV)	Amplitude	A : 11.0 ± 7.5 P : 7.4 ± 6.0	A : 9.5 ± 5.4 P : 5.5 ± 3.3	Non significatif
	Effet sur la sensibilité à la douleur	Douleur faible	A : 1.49 ± 0.94 P : 2.3 ± 4.2	A : 3.22 ± 8.7 P : 3.05 ± 4.2	Non significatif
		Douleur modérée	A : 3.3 ± 6.5 P : 1.5 ± 4.2	A : 3.9 ± 9.8 P : 1.8 ± 4.2	Non significatif
		Douleur sévère	A : 1.4 ± 4.5 P : 1.1 ± 2.9	A : 0.10 ± 0.31 P : 0.92 ± 3.1	Non significatif
	Consommation d'analgésiques par semaine toute molécule confondue	Nombre de comprimés utilisés en moyenne par semaine	Semaine 1 : P : 2.8 ± 5.9 A : 2.1 ± 5.3	Semaine 6 : P : 1.4 ± 2.6 A : 0.89 ± 2.0	Non significatif
	Effets secondaires (Pourcentage de patients concernés)	Tout effet indésirable confondu	Pas de résultat disponible		Non significatif
Spécifique amitriptyline		Sécheresse buccale P : 17% A : 54%	Somnolence : P : 27% A : 62%	Sécheresse buccale : p < 0.05 Somnolence : p < 0.05	

3. Pfaffenrath 1994 : Efficacy and tolerability of amitriptylinexide in the treatment of chronic tension-type headache: a multi-centre controlled study

a) Caractéristiques de l'étude :

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles
Participants	N = 197, adultes âgés de 18 à 65 ans, non hospitalisés présentant les critères IHS de la céphalée de tension chronique. Les céphalées devaient être présentes plus de 15 jours par mois pendant au moins 6 mois
Interventions	Les participants ont reçu soit : - Un comprimé de 30 mg d'amitriptylinexide, soit - Un comprimé de 25 mg d'amitriptyline soit - Un comprimé de placebo Une augmentation progressive est réalisée de telle sorte qu'à partir de 9 semaines les participants passent à 2 comprimés et à partir de la 13 -ème semaine à 3 comprimés. Si des effets secondaires apparaissent la dose est diminuée au pallier inférieur toléré.
Outcomes	<u>Objectifs principaux :</u> Fréquence des céphalées (nb jours/mois) Durée des céphalées (nb d'heures de céphalées/jours x mois) Fréquence des céphalées x Durée des céphalées L'intensité (VAS de 0 à 8) Enregistrement des effets indésirables avec le questionnaire FSUCL <u>Objectifs secondaires :</u> L'auto évaluation des patients par le "Self Report Symptom Inventory 90 items-R"
Notes	Critères d'exclusion : l'association à une migraine, la participation à une étude au cours des 3 derniers mois, Une mauvaise observance suspectée, les femmes enceintes ou allaitantes, la prise d'un simple analgésique, d'analgésiques mixtes, de tartrate d'ergotamine ou de tartrate de di-hydroergotamine, d'acide acétylsalicylique et/ou de paracétamol ou de codéine plus de 10 jours par mois, la prise d'autres antidépresseurs, neuroleptiques, tranquillisants (de manière régulière), de traitement prophylactique des maux de tête établie (par exemple, bêta-bloquants ou antagonistes du calcium) moins de 3 mois avant le début de la phase

	initiale, les médicaments pour le traitement prophylactique des troubles affectifs bipolaires (lithium et carbamazépine), des conditions médicales où le médicament prescrit pourrait entraîner des maux de tête comme effet secondaire. Les inhibiteurs de la MAO devaient avoir été arrêtés dans les 4 semaines précédant le début de l'étude. Étude multicentrique : 7 centres dans 3 pays Durée 24 semaines En ITT Patients randomisés : Amitritpylinoxide (AO) n = 66 , Amitriptyline (A) n = 67, Placebo (P) n = 64 Nombre de patients perdus de vue : 62 Résultats non significatifs contre placebo
--	---

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« The randomization was done in each center in blocks of 15 patients, that is , in each centre, series of 15 patients were treated at random with AO, AM or PL » « This was a three-armed, double-blind and placebo-controlled parallel group study in seven centres and three countries » « Dans certaines situations, un jugement peut être porté pour répondre « probablement non » ou probablement oui ». Par exemple, dans le contexte d'un grand essai par une unité d'essai clinique expérimentée [...] est susceptible d'entraîner une réponse plutôt « probablement oui » que « aucune information ».	PY
1.2 La séquence d'allocation est-elle dissimulée jusqu'à l'inscription et la répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-	« The individual groups were comparable with regard to sex mean age and duration of illness »	PN

elles un problème au niveau du processus de randomisation ?		
Evaluation du risque de biais		Quelques inquiétudes

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Pas de différence significative en ce qui concerne les effets indésirables pour les 3 groupes ni en général ($P_{AM-PL} = 0.5188$, $P_{AO-PL} = 1.000$) ni en effets indésirables spécifiques Mention d'étude en double aveugle et utilisation d'un placebo, 3 groupes de traitements	PN
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PN
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		NC
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		NC
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		NC
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITT	PY
2.7 Si N/PN/NI à 2.6 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		NC
Evaluation du risque de biais		Risque faible

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	62 patients / 211 randomisés => 31.4 % p _{dv}	PN
3.2 Existe-t-il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Utilisation par les auteurs de « last visit carried forward method »	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	1/4 à 1/3 de perdus de vue pour un seuil de 5% de p _{dv} admis	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ? (	“The reasons most commonly given for discontinuation were: - poor compliance - lack of efficacy - side effects”	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Double aveugle, pas de différence significative en ce qui concerne les effets indésirables, et autoévaluation	PN
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?		Non concerné
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?		Non concerné
Evaluation du risque de biais		Risque faible

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas de base de données d'enregistrement de protocole lors de la réalisation de l'étude	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Absence d'informations concernant la dissimulation de la séquence d'allocation
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque faible	Double aveugle a priori maintenu Pas d'écart par rapport à l'intervention prévue Analyse appropriée en ITT
Biais dû aux données manquantes	Risque élevé	31.4% de perdus de vue Utilisation de la méthode dite "de la dernière visite rapportée" pour la gestion des données manquantes Causes d'abandon a priori dépendantes de la vraie valeur
Biais dû à des erreurs de mesure dans les résultats	Risque faible	Critère de jugement subjectif mesuré par les participants Double aveugle maintenu
Biais dans la sélection des résultats	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement.
Autres biais		Pas d'information sur les conflits d'intérêts Méthode d'ajustement du risque alpha par Bonferroni
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Pfaffenrath année : 1994						
Etude randomisée parallèle en 3 bras : Amitriptylinoxide , Amitriptyline et Placebo N = 197, adultes de 18 à 65 ans Patients randomisés : Amitriptylinoxide (AO) n = 66 , Amitriptyline (A) n = 67, Placebo (P) n = 64 Analyse en ITT Valeur de significativité : Amitriptyline contre placebo : p ; Amitriptylinoxide contre placebo : p* Abréviations : Non significatif : NS, semaine S						
	Critère de jugement	Mesure du critère	Valeur rapportée	Résultats		Significatif ou non vs placebo
Critères de jugement principaux	Fréquence des céphalées	Nombre de jours / mois = 4 semaines	Moyenne ± écart type	Base: AO: 17 ± 8 A: 16 ± 8 P : 15 ± 8	4 dernières S: AO : 16 ± 9 A : 15 ± 10 P : 16 ± 9	Non significatif
	Durée des céphalées	Nombre d'heures par jour / mois		AO : 11 ± 5h A : 10 ± 7h P : 9 ± 6h	AO : 7 ± 6 h A : 7 ± 7h P : 7 ± 6h	Non significatif
	Intensité	VAS de 0 à 8		AO : 3.9 ± 1.6 A : 3.7 ± 1.9 P : 3.4 ± 1.5	AO : 2.8 ± 1.8 A: 2.8 ± 2.0 P: 1.7 ± 2.0	Non significatif
		Nombre de patients ayant une diminution de l'intensité	%	Diminution entre baseline et dernier mois AO : 71.2% A : 50.7 % P : 57.8%	Diminution > 50% Entre baseline et dernier mois AO : 31.8% A : 26.9% P : 26%	Diminution > 50% p et p* : NS Diminution : p et p* non fournis
	Fréquence x Durée des céphalées	Score	Moyenne et Écart type	AO : 299 ± 152 A : 280 ± 184 P : 250 ± 158	AO : 209 ± 167 A : 207 ± 194	p (A vs P) NS p* (AO vs P) NS

					P : 208 ± 171	
		Nombre de patients ayant une diminution du score	%	Diminution entre base et dernier mois AO : 71.2 % A : 64.2% P : 56.3 %	Diminution > 50 % entre base et dernier mois AO : 39.4 % A : 25.4 % P : 20.6%	Diminution : p et p* non fournis
	Amélioration des patients après 12 semaines de traitement	Durée x fréquence	%	AO : 51.5 %* A : 41.8 % P : 28.1 %		p : NS p* (AO vs P) < 0.01 Fisher's exact test
		Durée x fréquence x intensité		AO : 36.4 % A : 28.4 % P : 20.3 %		p et p* NS Rq : la tendance d'amélioration de l'intensité est significative pour l'intensité
		Intensité		AO : 43.9 % A : 32.8 % P : 28.1 %		
		Ressenti global positif		AO : 39.4 % A : 79.9 % P : 32.8 %		
		Ressenti positif sur la durée		AO : 36.4% A : 17.9% P : 23.4%		
		Ressenti positif sur l'intensité		AO : 36.4% A : 22.4% P : 28.1%		
	Réduction ≥50 % de durée x fréquence Et ≥ 50 % dans l'intensité des céphalées	Nombre de patients avec amélioration 50% - durée x fréquence ET - intensité douleur	AO : 30.3% A : 22.4% P : 21.9 %		Non significatif	

	Enregistre- ment des effets indésirabl es	Effets indésirables (EI) :		AO : 78.8 % A: 88.1 % P: 79.1 %		Non significatif
	Question- naire FSUCL	EI modérés à sévères :		AO: 72.7 A : 73.1 P : 57.8		
		EI en lien avec les traitements:		AO : 62.1 % A : 68.7 % P : 46.9 %		
		Abandons pour cause d’EI :		AO : 3% A : 17.9% P : 10.9%		
Critère de jugement secondair e	L'auto- évaluation des patients	"Self Report Symptom Inventory 90 items-R"	<u>Succès :</u>	“After 3 months AO was regarded as more effective than both comparison substances” “there was an insignificant difference between A and P”		Non significatif
		En pour cent	<u>Pire</u>	AO : 5 % “ and slightly greater in both comparison groups”		
			<u>pas d’amélior ation</u>	AO : 24 % A : 23.9 % P : 34.4%		
			Score de Dépression	Moyenne et écart type	<u>Baseline</u> AO :0.58 ± 0.52 A: 0.76 ± 0.68 P: 0.60 ± 0.63	<u>Dernière visite</u> AO : 0.45 ± 0.50 A: 0.70 ± 0.70 P : 0.51 ± 0.62
Autres	Consomm ation d’antalgiqu es	Patients n’ayant pas eu recours à des antalgiques	Pourcent -age	AO : 31.0% A : 25.6 % P : 12%		Non significatif

4. Brendtsen 1996: A non-selective (amitriptyline), but not a selective citalopram), serotonin reuptake inhibitor is effective in the prophylactic treatment of chronic tension-type headache.

a) Caractéristiques de l'étude

Méthode	Essai prospectif, randomisé, en double aveugle, contrôlés par un placebo, en crossover avec 3 bras.
Participants	N = 40 Adultes âgés de 18 à 65 ans présentant une céphalée de tension chronique diagnostiquée selon les critères de l'International Headache Society.
Interventions	Les participants ont reçu les 3 traitements suivants consécutivement dans un ordre variable : - De l'Amitriptyline à la dose de 25 mg par jour initialement augmentée à 50 mg la deuxième semaine puis à 75 mg par jour à partir de la 3ème semaine. - Du Citalopram à la dose de 20 mg par jour - Un Placebo
Outcomes	<u>Critère de jugement principal :</u> - Aire sous la courbe des céphalées calculée comme la somme des enregistrements quotidiens de la durée de la céphalée x l'intensité de la céphalée. <u>Critères de jugement secondaire :</u> - La durée des céphalées au cours des quatre dernières semaines de chaque période de traitement - L'intensité moyenne des céphalées par jour de céphalées - La fréquence des céphalées au cours des quatre dernières semaines de chaque période de traitement - Le nombre de doses d'analgésiques, équivalent à 500 mg d'aspirine, pris au cours des quatre dernières semaines de chaque période de traitement - Le nombre d'effets indésirables au cours de chaque période de traitement
Notes	Critères d'exclusion : La participation antérieure à un essai clinique, un épisode de migraine plus d'un jour par mois, les maladies somatiques ou psychiatriques graves y compris la dépression (score de dépression de Hamilton supérieur ou égal à 17), le mésusage d'analgésiques simples (correspondant à plus de 2 g d'aspirine par jour), une prise régulière d'opiacés ou de benzodiazépines et des traitements antérieurs par des antidépresseurs.

	Les patients ont été informés que l'étude inclus des périodes de placebo, mais aucune autre information sur la conception de l'étude n'a été donnée. Les femmes en âge de procréer devaient utiliser des mesures contraceptives adéquates tout au long de l'étude. Etude monocentrique Période initiale "run in period" : 4 semaines Période de "wash-out" : 2 Semaines Période de traitement : 8 Semaines Durée totale 32 semaines Analyse en ITTm Nombre de patients analysés : 34 Nombre de perdus de vue : 7 Résultat significatif pour l'amitriptyline versus placebo : Réduction de 30 % de la courbe AUC, diminution significative de la durée et de la fréquence des céphalées.
--	---

b) Tableau RoB 2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« randomly allocated » « randomisation was done in blocks of 6 patients »	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	<i>"None of the clinical characteristics presented in table 1 differed between the 34 patients who completed the study and the six drop outs (P = 0.32 to 0.98) or between the groups of patients allocated to the six different treatment sequences (P = 0.20 to 0.80)"</i>	PN
Evaluation du risque de biais		Quelques inquiétudes

Biais dû à l'effet « carry-over »

Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	“Each patient was randomly allocated to one of the six treatment sequences a-f “	PY
S2. Si N, PN, NI à question S.1 : Est-ce que la période d’effet est prise en compte dans l’analyse ?		
S3. Est-ce qu’il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	« A possible carry-over effect was tested [...]by means of the Mann-Whitney two tailed test. » « There was no carry-over effect (P = 0.06 to 0.39) »	Y
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l’affectation à l’intervention)

Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l’intervention qui leur étaient assignée pendant l’essai ?	« All tablets were identical look and taste »	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l’intervention des patients pendant l’essai ?	« Amitriptyline induced significantly more side effects than both citalopram and placebo (P < 0.001), whereas there was no difference between citalopram and placebo (P> 0.2 ; table 3). The difference between amitriptyline and placebo was due to a higher number of patients complaining of dry mouth (p<0.001) and of drowsiness (p< 0.001)during the former treatment. » Follow up visits were performed at four-week intervals (fig 1). At each visit, [...]side effects reported by patients were recorded »	PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l’intervention envisagée en raison du contexte de l’essai ?		PN

2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITTm « patients who drop out of the study were excluded from the analyses » donc 6 patients exclus sur 40 soit 15% d'exclusions.	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	Plus de 5% de perdus de vue avec analyse en ITTm	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	6 exclus car perdus de vue Donc 15% de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas de test réalisé	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	« The reasons for drop out were side effects (one patient on amitriptyline [...]), pregnancy (one patient on placebo), and lack of effect (two patients on placebo and two on citalopram) » Le manque d'effet et effets indésirables sont liés au traitement soit 5 perdus de vue exclus en lien avec les résultats sur 40 patients => 12,5% données manquantes en lien avec résultat	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	Oui 12.5% données manquantes dépendantes de leur vraie valeur.	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Evaluateur : médecin et/ou patient Significativement plus d'effets indésirables de l'amitriptyline que dans le placebo y compris plus d'effets indésirables spécifiques de l'amitriptyline.	PY
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Oui, critères de jugements subjectifs	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Article de 1996 Protocole a priori non enregistré dans une base d'enregistrement	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	Tous les résultats sont rapportés	PN
5.3 ... Plusieurs analyses éligibles des données.		PN
5.4 : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Peu d'informations sur la randomisation Absence de différences observées entre la population de base
Biais dû à l'effet « carry-over »	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle a priori levé Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	15% de données manquantes Données manquantes a priori dépendantes de la vraie valeur
Biais dû à des erreurs de mesure dans les résultats	Risque élevé	Double aveugle a priori levé Evaluateur : patient Critères de jugements subjectifs
Biais dans la sélection des résultats	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole a priori non enregistré préalablement Absence de sélection des résultats Une seule analyse réalisée Résultats de toutes les périodes disponibles
Autres biais		Possible biais de confusion : absence d'exclusion entre autres des patients ayant déjà pris les molécules testées, des patients sous traitement psychotrope autre que les benzodiazépines, des patients présentant des addictions. Absence d'informations concernant les conflits d'intérêts Méthode d'ajustement du risque alpha par Bonferroni
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur :Brentsen Année: 1996						
	Etude en crossover : 3 bras avec Amitriptyline/Citalopram/Placebo Analyse en ITTm Nombre de participants initialement : N = 40 Nombre de patients analysés : N = 34 Nombre de perdus de vue : N = 6 Durée de l'étude : 32 semaines Période initiale "run in period" : 4 semaines Période de "wash-out" : 2 Semaines Période de traitement : 8 Semaines Multiplication des critères de jugements avec ajustement "To correct for multiple testing the P values were multiplied according to the method of Bonferroni."					
	Critère de jugement	Mesure du critère	Valeur rapportée	Amitriptyline	Placebo	Significatif ou non versus placebo
Critère de jugement primaire : (Mesuré les 4 dernières semaines de chaque période de traitement)	Réduction courbe AUC céphalées	Somme des enregistrements quotidiens de la durée x l'intensité	Moyennes	616 (129)	877 (171)	Significatif (selon la période $p < 0,02$, ou $p < 0,008$ ou $p < 0,001$) Friedman's test et si différence significative Wilcoxon's paired rank

			% de diminution AUC	« The AUC was 30% lower on amitriptyline than on placebo »		P = 0.002	
	Efficacy of amitriptyline ? = (1-AUC _{amitriptyline} / AUC _{placebo}) x 100% => amélioration > 30%		Nombre de patients	27/34			
Critère de jugement secondaire (Mesurés sur les 4 dernières semaines de chaque période de traitement)	Durée des Céphalées	Nombre d'heures / 4 semaines	Heures	151 (24)	184 (27)	P = 0.01	
	Intensité des Céphalées	(EVA 0->10 mm)	Moyenne	3.8 (0.3)	3.9 (0.3)	P = 0.12	
	Fréquence des Céphalées	Nombre de journées sur 4S	Journées /4S	18.6 (1.6)	21.7 (1.3)	P = 0.01	
	Nombre d'analgésiques pris (équivalent 500 mg aspirine)		Doses / 4 semaines	25.3 (4.2)	33.8 (5.2)	P = 0.02	
				Amitriptyline	Citalopram	Placebo	Significatif vs placebo
	Effets secondaires	Généraux	Nombre de patients	33	15	15	P < 0.001
		Sècheresse buccale		30	5	3	P < 0.001
		Somnolence		21	7	6	P < 0.001

5. Diamond 1998 : Chronic tension headache--treated with amitriptyline a double-blind study

a) Caractéristiques de l'étude :

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles
Participants	N = 90, adultes de 18 à 68 ans, présentant de l'anxiété ou une dépression et se plaignant de céphalées. Groupe Amitriptyline 10 mg N=30, Groupe Amitriptyline 25mg N=30, Groupe Placebo N = 30
Interventions	Les participants ont reçu au hasard l'une des 3 molécules ci-dessous - 10 mg d'amitriptyline, : - 25 mg d'amitriptyline - un placebo Initialement 2 comprimés ont été administrés au coucher, les doses ont pu être augmentées à partir de la première semaine jusqu'à atteindre au maximum 6 comprimés par jour à la 4ème semaine.
Outcomes	Absence de différenciation entre critère primaire et secondaire <u>8 symptômes cibles ont été évalués :</u> L'anxiété, l'humeur dépressive, les plaintes psychosomatiques, l'insomnie, les céphalées, l'apathie, la fatigue, la baisse de l'efficacité au travail Ces symptômes étaient notés de 1 à 4 avec 1= aucun, 2= léger, 3= modéré, 4=sévère. La réponse globale hebdomadaire au traitement par le patient et l'évaluation globale faite par le médecin étaient notées de 1 à 4 également, avec 1 = aucune, 2 = égale, 3 = bonne, 4 = excellente). Le poids des patients, la tension artérielle, le pouls étaient enregistrés à chaque visite. Une numération sanguine, une analyse d'urines, l'azote d'urée dans le sang (BUN) et les transaminases étaient mesurés initialement et à la fin de l'étude.
Notes	Critères d'exclusion : Les patients ne présentant pas de céphalées définies comme légères/modérées/sévères et les patients ne présentant aucun des symptômes ciblés n'étaient pas inclus dans l'étude. Etude monocentrique, Durée 4 semaines Analyse en ITT, Participants analysés : N=55 AT 25 mg : 18 patients ; AT 10 mg : 21 patients ; placebo 16 patients 5 patients exclus avant le début de l'essai et 27 abandons Résultat significatif pour l'amitriptyline 10 mg versus placebo dans le traitement des céphalées à 1 semaine uniquement

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?		NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	Données de population disponibles (âge, sexe, céphalées, anxiété, dépression) en population finale (données sur 85 patients et non sur 90)	NI
Evaluation du risque de biais		Quelques inquiétudes

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	<i>"The size, shape and color of all three tablets were identical."</i> Tableau VI " <i>side effects : total patients with side effects</i> <i>Placebo : 1 sur 29 patients</i> <i>Amitriptyline 10 mg : 8* sur 28 patients</i> <i>Amitriptyline 25 mg : 8* sur 28 patients</i> <i>* Statistically significantly greater than placebo "</i>	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non Concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention		Non Concerné

envisagée étaient-ils équilibrés entre les groupes ?		
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITT pour « physician global évaluation » car réalisé sur 85 patients Analyse en ITTm pour « headache » à semaine 4 : Amitriptyline 10 mg: 21 patients sur 28 Amitriptyline 25 mg: 18 patients sur 28 Placebo : 16 patients sur 29 Donc 55 patients analysés en S4 sur les 85 initiaux soit exclusion de 30 patients : 35% d'exclu d'analyses Donc ITTm inappropriée	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	<i>“Of these 85 patients, twenty seven (32%) did not complete the entire four-week course of study for various reasons”</i>	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas de méthode de gestion de données faite pour l'ITTm Pour l'analyse faite en ITT pour l'évaluation globale le médecin : valeur attribuée aux données manquantes “pas d'amélioration”	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Pas d'information sur les causes de drop out mais hypothèses des auteurs d'abandon pour cause d'inefficacité et d'effets indésirables	PY

	<i>“The severity of the anti-cholinergic side effects could have as easily been the reason for these patients not returning or dropping from the study. This would not necessarily, then, be related to the response of this drug as an active therapeutic agent”</i>	
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l’absence de résultat dépende de sa vraie valeur ?	Plus de 32% de données manquantes	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d’un groupe d’intervention à l’autre ?	Le nombre de visites par patient et l’espacement entre les visites peut varier selon les patients « <i>Since in some cases the times when the patients were seen varied, if they were seen at week 2 or 3 this was counted as week 2 ; if they were seen at week 4 or 5, this was counted as week 4. If any patient had two measurements in an interval the measurement were averaged.</i> »	PY
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l’intervention reçue par les participants à l’étude ?	Double aveugle non maintenu Evaluateurs : médecin et/ou patient	PY
4.4 Si Y / PY / NI à 4.3 : L’évaluation du résultat aurait-elle pu être influencée par la connaissance de l’intervention reçue ?	Critères de jugements subjectifs	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l’évaluation des résultats ait été influencée par la connaissance de l’intervention reçue ?	Pas de critères permettant de statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Etude ancienne, pas de base d'enregistrement de protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	« The weight, blood pressure and pulse data recorded initially and at the end of weeks 1,2and 4 were tranformed to algorithms to facilitate an analysis of the data according to a more normal distribution. Both amitriptyline treatment groups showed pulse rates with statistically significantly larger increases from pretreatment than placebo. This was seen in the amitriptyline 10 mg groups at weeks and 2?? and in the amitriptyline 25 mg group at week 1. There we differences among groups in weight and / or blood pressure measurements. » Ne rapporte pas les résultats de l'ensemble des mesures pour la fréquence cardiaque.	PY
5.3 ... Plusieurs analyses éligibles des données.	<i>"If these nine drop-outs are included in the evaluation of the results as having poor response we arrive at the statistics as presented here »</i> <i>" We feel that these drop outs should not have been used as part of the statistical evaluation and if you remove this number from that particular dose group the small number remaining are not enough to allow you to show statistical significance for this particular group"</i>	PY
Evaluation du risque de biais		Risque élevé

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Méthodologie non décrite. Données de la population disponibles uniquement pour la population finale.
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Doute important concernant le maintien du double aveugle (effets indésirables du traitement majeurs dans les groupes Amitriptyline). Analyse inappropriée tantôt en ITT tantôt en ITTm.
Biais dû aux données manquantes	Risque élevé	32% d'abandon. Probabilité importante que le manque de résultat dépende de sa vraie valeur.
Biais dû à la mesure de l'outcome	Risque élevé	Doute important concernant le maintien du double aveugle. Critère de jugement subjectif. La connaissance du statut de l'intervention aurait pu influencer l'évaluation des résultats
Biais dû à la sélection des résultats rapportés	Risque élevé	Certains résultats annoncés par le protocole non entièrement rapportés. Absence de temporalité du protocole Réalisation de plusieurs analyses avec sélection de l'analyse la plus appropriée.
Autres biais		Possible risque de biais de confusion : absence d'exclusion des patients sous psychotropes ou autres médicaments. Etude sur une population exclusivement dépressive. Incompréhension du protocole : augmentation des doses d'amitriptyline sans information sur les doses finales administrées. Pas d'information sur les conflits d'intérêts. Pas de méthode d'ajustement du risque alpha rapportée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur: **Diamond** Année : **1998**

Essai randomisé en parallèle 3 bras d'intervention : Amitriptyline 10 mg, Amitriptyline 25 mg , Placebo Analyse en ITT pour l'évaluation globale par le médecin et en ITTm pour les autres critères de jugement AT 10mg n = 28 (5 drop out), AT 25 mg n = 28 (9 drop out), placebo n = 29 (13 drop out)							
	Critère de jugement	Mesure du critère	Valeur rapportée	Amitriptyline		Placebo	Significatif ou non versus placebo
				10mg	25mg		
Pas de classement entre critère de jugement primaire ou secondaire	Anxiété	Autoévaluation échelle	Moyenne des scores pré et post traitement de S1, S2, S4	S1 0.8	0.4	0.2	p < 0.05
				S2 1.0	0.9	0.6	
				S4 1.0	0.7	0.7	
	Humeur dépressive	1= aucune 2 = faible 3= modéré		S1 1.3	0.8	0.6	
				S2 1.9	1.4	1.2	
				S4 2.3	1.6	1.7	
	Plaintes psychosomatiques	4= important		S1 1.2	0.7	0.5	
				S2 1.8	1.5	1.1	
	Insomnie		Rq : tout patient qui a, à l'inclusion dans l'étude, une valeur de 1 (aucun) pour un des symptômes de la colonne de gauche, est exclu	S4 2.2	1.6	1.3	
				S1 1.1	0.9	0.6	
				S2 1.6	1.6	1.1	
				S4 2.1	1.8	1.2	
	Céphalées			S1 1.5	0.8	0.6	
				S2 1.7	1.5	1.5	
				S4 2.1	1.5	1.3	
	Apathie			S1 1.0	0.4	0.2	
				S2 1.4	0.8	0.6	
				S4 1.5	0.9	0.3	
	Fatigue			S1 0.8	0.2	0.2	
				S2 1.4	0.6	0.7	
				S4 1.5	0.7	0.5	
	Diminution de l'efficacité au travail			S1 0.8	0.3	0.2	
				S2 1.4	0.6	0.6	
				S4 1.5	0.8	0.5	

		de l'analyse de l'évolution de ce symptôme.					
	Evaluation globale par le patient	Autoévaluation	Score moyen semaine 1, 2, 4	S1 2.43	1.75	1.64	p < 0.05
		1 = aucune		S2 2.98	2.5	2.19	
		2 = légère		S4 3.38	2.61	2.44	
	Evaluation globale par le médecin	3 = bonne					
		4 = excellente					
	Evaluation globale par le médecin	Echelle amélioration 1 = aucune 2 = légère 3 = bonne 4 = excellente	Moyenne	3.07	2.44	2.04	p < 0.01
	Nombre de patients avec effets Indésirables	Somme des patients ayant au moins 1 EI		8/28	8/28	1/29	p < 0.05
		Remarque :	Somnolence Bouche sèche	3 3	5 5	0 0	

6. Brendtsen 2000: Amitriptyline reduces myofascial tenderness in patients with chronic tension-type headache

a) Caractéristiques de l'étude :

Méthode	Essai prospectif, randomisé, en double aveugle, contrôlé par un placebo, en crossover avec 3 bras.
Participants	N= 40 Adultes âgés de 18 à 65 ans présentant une céphalée de tension chronique diagnostiquée selon les critères de l'International Headache Society.
Interventions	Les participants ont reçu les 3 traitements suivants consécutivement dans un ordre variable : - De l'Amitriptyline à la dose de 25 mg par jour initialement augmentée à 50 mg la deuxième semaine puis à 75 mg par jour à partir de la 3ème semaine. - Du Citalopram à la dose de 20 mg par jour - Un Placebo
Outcomes	<u>Critères de jugement principal :</u> “ l'amitriptyline inhibe principalement la douleur musculaire comme en témoigne une réduction de la sensibilité myofasciale ou si elle réduit la sensibilité à la douleur en général comme en témoigne l'augmentation des seuils de douleur mécanique électrique ou extracrânienne “ - La sensibilité myofasciale - Le seuil de douleur à la pression - Le seuil de douleur électrique <u>Critères de jugement secondaire :</u> - Aire sous la courbe des céphalées calculée comme la somme des enregistrements quotidiens de la durée de la céphalée x l'intensité de la céphalée - L'intensité moyenne des céphalées par jour de céphalées
Notes	Critères d'exclusion : la survenue d'épisode migraineux de plus d'une 1 journée par mois, les maladies somatiques ou psychiatriques graves incluant la dépression (Score de dépression d'Hamilton supérieur ou égal à 17), la prise de médicaments prophylactiques pour des céphalées de tension, l'abus d'analgésiques simples (correspondant à plus de 2 g d'aspirine/jour), la prise régulière d'opiacés ou de benzodiazépines, un traitement antérieur par anti-dépresseurs.

	Etude monocentrique Période initiale “run in period” : 4 semaines Période de “wash-out” : 2 semaines Période de traitement : 8 semaines Analyse en ITTm Nombre de patients analysés : 34 Nombre de perdus de vue : 6 Résultats significatifs pour l’amitriptyline versus placebo pour la réduction de la courbe AUC et pour l’intensité des céphalées.
--	--

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation réponses données dans la méthodologie de Bendtsen 1996 :		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d’allocation est-elle aléatoire ?	« randomly allocated » « randomisation was done in blocks of 6 patients »	NI
1.2 La séquence d’allocation est-elle dissimulée jusqu’à inscription et répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d’intervention suggéraient-elles un problème au niveau du processus de randomisation ?	None of the clinical characteristics presented in table 1 differed between the 34 patients who completed the study and the six drop outs (P = 0.32 to 0.98) or between the groups of patients allocated to the six different treatment sequences (P = 0.20 to 0.80)	PN
Evaluation du risque de biais		Quelques inquiétudes

Biais dû à l’effet « carry-over » réponses données dans la méthodologie de Bendtsen 1996 :		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	“Each patient was randomly allocated to one of the six treatment sequences a-f dans la méthodologie de Brentsen 1996.	PY

S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?		Non concerné
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	« A possible carry-over effect was tested [...] by means of the Mann-Whitney two tailed test. » « There was no carry-over effect (P = 0.06 to 0.39) »	Y
		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Bendtsen 2000 : « Although amitriptyline has clearly identifiable side effects, effective blinding was ensured by the rather complicated three way cross over designed » Mais même design que Bendtsen 1996 où les auteurs avaient trouvé 2 fois plus d'effets secondaires spécifiques pour l'amitriptyline que pour le citalopram et le placebo. Par ailleurs les effets secondaires ne sont pas quantifiés. Doute sur le maintien du double aveugle. Bendtsen 1996 : « All tablets were identical look and taste » « « Amitriptyline induced significantly more side effects than both citalopram and placebo (P < 0.001), whereas there was no difference between citalopram and placebo (P> 0.2 ; table 3). The difference between amitriptyline and placebo was due to a higher number of patients complaining of dry mouth (p<0.001) and of drowsiness (p< 0.001)during the former treatment. »	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY

	Follow up visits were performed at four-week intervals (fig 1). At each visit, [...] side effects reported by patients were recorded »	
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?	“The recordings of electrical pain sensitivity were not completed in another 6 patients due to technical problems “	PY
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?	Absence d'informations fournies	NI
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?	Absence d'informations fournies	NI
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Probable analyse en ITTm car Bendtsen 1996 en ITTm et que dans la légende de la figure 2 « (a) Total Tenderness Score (TTS) in 33 patients with chronic tension-type headache during treatment with amitriptyline 75 mg/day, citalopram 20 mg/day or placebo » donc 7 patients exclus sur 40 soit 17.5% d'exclusion	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	17.5% d'exclus	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	Pour le critère de jugement principal, le nombre de données manquantes varie de - 7 perdus de vue pour la mesure « myofascial tenderness » sur 40 - 13 données manquantes (7 pdv + 2 pb matériel + 4 refus de participation) pour la mesure de « pain sensitivity in general » sur 40 soit 32.5% Les critères de jugements secondaires sont concernés par 7 perdus de vue sur 40	PN
3.2 Existe il des preuves que les résultats n'aient pas été	Pas de test réalisé	PN

biaisés par les données manquantes ?		
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	32.5 % de données manquantes	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	« Six patients dropped out of the trial for various reasons (4) unrelated to the present study, while one patient did not want to participate in the present study during the trial. The recordings of electrical pain sensibility were not completed in another six patients [...] due to technical problems (n = 2) and unwillingness to be subjected to electrical current (n = 4) » Donc a priori pas en lien avec la vraie valeur.	PN
Evaluation du risque de biais		Quelques inquiétudes

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	L'évaluateur est le patient (autoévaluation des critères). Or, bien que les auteurs affirment que du fait de la complexité de leur design d'étude (crossover 3 bras) le double aveugle a pu être conservé, cependant on a vu dans l'étude de 1996 du même auteur que ce n'était pas le cas. Donc il y a un risque non négligeable que les patients aient été au courant de l'intervention reçus et comme c'est une autoévaluation, les évaluateurs sont au courant de l'évaluation reçue.	PY
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée	Oui, critères de jugements subjectifs	PY

par la connaissance de l'intervention reçue ?		
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information sur un enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	Tous les résultats sont rapportés	PN
5.3 ... Plusieurs analyses éligibles des données.		PN
5.4 : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Quelques inquiétudes

c) **Résumé des domaines et jugement global :**

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Peu d'informations sur la randomisation Absence de différences observées dans la population de base
Biais dû à l'effet carry-over	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle a priori levé Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Quelques inquiétudes	32.5 % de données manquantes Données manquantes a priori non dépendantes de la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Double aveugle a priori levé Evaluateur : patient Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole a priori non enregistré préalablement Absence de sélection des résultats Une seule analyse réalisée Résultats de toutes les périodes disponibles
Autres biais		Possible biais de confusion : absence d'exclusion entre autres des patients ayant déjà pris les molécules testées, absence d'exclusion des patients sous autres médicaments psychotropes que les benzodiazépines, des patients présentant des addictions. Absence d'informations concernant les conflits d'intérêts Méthode d'ajustement par Bonferroni réalisée en 1996 donc probablement faite également
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Brendtsen année : 2000						
Etude randomisée en crossover 3 bras d'intervention : Amitriptyline, Citalopram, Placebo Analyse en ITTm Participants : N = 40						
	Critère de jugement	Mesure du critère	Valeur rapportée	Amytriptyline Mmol/L	Placebo	Significatif ou non versus placebo
Critère de jugement secondaire	Réduction courbe AUC céphalées amitriptyline vs placebo (EVA 0->10cm)	Wilcoxon's paired rank : Différence entre AUC céphalées lors du traitement par amitriptyline et par placebo	Valeur médiane + quartiles en parenthèse	« Amitriptyline reduced AUC by 30% compared with placebo » Pas d'autre information dans l'article		« highly significant » Pas d'autre information dans l'article
	Intensité des céphalées	(EVA 0->100 mm)		10 (0 – 22)	16 (0 – 37)	P = 0.04
Critère de jugements principaux	Sensibilité myofaciale (Total Tenderness Score : somme score intensité 0-3 sur 18 points, max 48)	Patients n=33		8 (3-12)	11 (5-16)	P = 0.01
		Patients répondeurs n = 19		3 (2-11)	10 (2-13)	P = 0.005
		Patients non répondeurs n = 14		15 (8-25)	13.5 (7-25)	P = 0.36
		Comparaison TTS répondeurs vs non répondeurs		« responders had significantly lower TTS than non responders »	non concerné	P = 0.003
	Sensibilité globale à la douleur (kPa)	PPDT doigt		432 (336-606) kPa	absence donnée	P = 0.8
		PPTO doigt		918 (624 – 1138)	808 (712 – 1102)	P = 0.28

		PPDT R. temporale		286 (218- 398)	278 (212 – 398)	P = 0.21
		PPTO R temporale		544 (406 – 782)	531 (405- 729)	P = 0.31
		REPT (mA) (EPDT – ESDT)		4.0 (2.1 – 6.0)	3.8 (2.0- 5.3)	P = 0.73
Critère de jugement secondaire «clarify the relative importance of serotonin reuptake inhibition for analgesic effect of amitriptyline by investigating whether the highly selective serotonin reuptake inhibitor citalopram affects the above mentioned parameters in patients with chronic tension type headache»				Comparaison citalopram vs placebo, donc résultats non rapportés dans ce tableau		

7. ASHINA et BRENDTSEN 2004 : Analgesic effect of amitriptyline in chronic tension-type headache is not directly related to serotonin reuptake inhibition.

a) Caractéristiques de l'étude :

Méthode	Essai prospectif, randomisé, en double aveugle, contrôlé par un placebo, en crossover avec 3 bras.
Participants	N= 40 Adultes âgés de 18 à 65 ans présentant une céphalée de tension chronique diagnostiquée selon les critères de l'International Headache Society.
Interventions	Les participants ont reçu les 3 traitements suivants consécutivement dans un ordre variable : - De l'Amitriptyline à la dose de 25 mg par jour initialement augmentée à 50 mg la deuxième semaine puis à 75 mg par jour à partir de la 3 -ème semaine. - Du Citalopram à la dose de 20 mg par jour - Un Placebo
Outcomes	<u>Critère de jugement principal :</u> - Le taux de 5 HT plaquettaire reflétant le contenu 5-HT dans les terminaisons nerveuses sérotoninergiques. <i>“The study investigates the importance of 5-HT reuptake inhibition for the analgesic effect of amitriptyline by comparison of 5 HT levels after treatment with amitriptyline and citalopram”</i>

	- La durée des céphalées au cours des quatre dernières semaines de chaque période de traitement - L'intensité moyenne des céphalées par jour de céphalées - Aire sous la courbe des céphalées calculée comme la somme des enregistrements quotidiens de la durée de la céphalée x l'intensité de la céphalée <u>Critères de jugement secondaires :</u> - Le taux de 5 HT plasmatique “ The secondary aim was to investigate plasma 5 HT levels simultaneously “ - Le nombre de doses d'analgésiques, équivalent à 500 mg d'aspirine, pris au cours des quatre dernières semaines de chaque période de traitement - Le nombre d'effets indésirables au cours de chaque période de traitement
Notes	Critères d'exclusion : la survenue d'épisode migraineux de plus d'une 1 journée par mois, les maladies somatiques ou psychiatriques graves incluant la dépression (Score de dépression d'Hamilton supérieur ou égal à 17), la prise de médicaments prophylactiques pour des céphalées de tension, l'abus d'analgésiques simples (correspondant à plus de 2 g d'aspirine/jour), la prise régulière d'opiacés ou de benzodiazépines, un traitement antérieur par anti-dépresseurs. Les aliments contenant de grandes quantités de sérotonine (par ex. bananes) ont été interdits pendant l'expérience. Les femmes en âge de procréer devaient utiliser des mesures contraceptives adéquates tout au long de l'étude. Etude monocentrique Période initiale “run in period” : 4 semaines Période de “wash-out” : 2 semaines Période de traitement : 8 semaines Analyse en ITTm a priori Nombre de patients analysés : 34 Nombre de perdus de vue : 6

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Réponses données dans la méthodologie de Bendtsen 1996		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	Non décrite dans cette étude mais « The study was designed as a double-blind, placebo-controlled, three-way crossover trial and has been previously described in details by Bendtsen et al, 1996 » Or dans l'étude de Bendtsen et al, 1996 il est dit « randomly allocated » « randomisation was done in blocks of 6 patients »	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	Pas de calcul de recherche de différence significative entre les caractéristiques de bases des individus des groupes initiaux Essai en crossover	NI
Evaluation du risque de biais		Quelques inquiétudes

Biais dû à l'effet « carry-over »		
Réponses données dans la méthodologie de Bendtsen 1996 : même nombre de patients, même nombre de perdus de vue, même schéma d'étude, mêmes molécules testées.		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	“Each patient was randomly allocated to one of the six treatment sequences a-f “ selon la méthodologie de Brentsen 1996	PY
S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?		
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	« A possible carry-over effect was tested [...]by means of the Mann-Whitney two tailed test. » « There was no carry-over effect (P = 0.06 to 0.39) »	Y
Evaluation du biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Pas d'information sur les effets indésirables alors que les patients les ont notés dans leurs carnets : « Throughout the study patients kept a headache diary with recordings of [...] side effects. » En utilisant le même design que Bendtsen 1996 où les auteurs avaient trouvés 2 fois plus d'effets indésirables spécifiques pour l'amitriptyline que pour le citalopram et le placebo. Il est donc attendu que cette étude qui reprend le même nombre de patient, avec le même nombre de perdus de vue et le même schéma et en testant les mêmes molécules donne les mêmes effets indésirables. <u>Bendtsen 1996 :</u> « All tablets were identical look and taste » « « Amitriptyline induced significantly more side effects than both citalopram and placebo ($P < 0.001$), whereas there was no difference between citalopram and placebo ($P > 0.2$; table 3). The difference between amitriptyline and placebo was due to a higher number of patients complaining of dry mouth ($p < 0.001$) and of drowsiness ($p < 0.001$) during the former treatment. » Follow up visits were performed at four-week intervals (fig 1). At each visit, [...] side effects reported by patients were recorded »	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention		Non concerné

envisagée étaient-ils équilibrés entre les groupes ?		
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Probable analyse en ITTm : au départ 40 patients et finalement uniquement 34 ont été analysés (tableau 1, figures 2 et 3). Donc 6 pdv ou exclus. Même méthodologie qu'en 1996 et le recours à une analyse en ITTm c'est-à-dire excluant de l'analyse uniquement des patients perdus de vue, alors on supposera que les auteurs ont fait de même cette fois ci. 40 patients randomisés, 6 pdv => exclusion de 15% de patients	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	15% d'exclus	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tous ou presque tous les participants randomisés ?	6 patients non inclus dans l'analyse finale soit 15% de données manquantes pour un seuil à 5%	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas de test réalisé	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Pas d'information sur les raisons de la non-analyse des données de ces patients. Probables perdus de vue, mais pas de raisons rapportées	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	15% de données manquantes sans fournir de raisons	NI
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Evaluateur : patient Doute sur le maintien du double aveugle	PY
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Oui, critères de jugements subjectifs	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement au préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
5.4 : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Quelques inquiétudes

c) **Résumé des domaines et jugement global :**

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Peu d'informations sur la randomisation Absence de données concernant la population de base
Biais dû à l'effet carry-over	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle a priori levé Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	15 % de données manquantes Absence d'information sur les causes de données manquantes
Biais dû à la mesure de l'outcome	Risque élevé	Double aveugle a priori levé Evaluateur : patient Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole a priori non enregistré préalablement Absence de sélection des résultats Une seule analyse réalisée Résultats de toutes les périodes disponibles
Autres biais		Possible biais de confusion : absence d'exclusion entre autres des patients ayant déjà pris les molécules testées, absence d'exclusion des patients sous autres médications psychotropes que les benzodiazépines, des patients présentant des addictions. Absence d'informations concernant les conflits d'intérêts Méthode d'ajustement faite en 1996 avec méthode de Bonferroni donc probablement réalisée ici aussi
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteurs : Ashina et Brendtsen Année : 2004						
Etude randomisée en crossover 3 bras d'intervention : Amitriptyline, Citalopram, Placebo Analyse en ITTm Participants : N = 40						
	Critère de jugement	Mesure du critère	Valeur rapportée	Amytriptyline Mmol/L	Placebo	Significatif ou non versus placebo
Critère de jugement primaire : « Investigate the importance of 5-HT reuptake inhibition for the analgesic effects »	Evolution du 5 HT dans les plaquettes lors du traitement par amitriptyline vs placebo	Wilcoxon's apired rank Différence entre AUC placebo et AUC amitriptyline sur le dosage du 5HT dans les plaquettes	Mediane + quartile	1.7 (1.2-2.4) x 10 ⁻¹⁸	3.5 (2.8 – 4.3) x 10 ⁻¹⁸	P < 0.001
	Calcul d'un coefficient de corrélation entre : - Evolution des céphalées sous Amitriptyline versus placebo - Evolution de la concentration 5HT dans les plaquettes sous amitriptylin	Sperman's test pour calculer les coefficients de corrélation		Différence AUC céphalées (amitriptyline vs placebo): 103 (10-437) Différence AUC 5HT dans les plaquettes groupe amitriptyline vs placebo : 1.7 (1.1-2.2) Coefficient de corrélation calculée : r = 0.12		P = 0.49

	e vs placebo				
	Céphalées : Verbal numérique rating scale : 0-> 10 AUC = duration x intensity of headache (somme des jours rapportés dans le journal)	Wilcoxon's paired rank : Différence entre AUC céphalées lors du traitement par amitriptyline et par placebo		308 (157 – 715)	441 (178-1408) P = 0.001
Critère de jugement secondaire « investigate plasma 5 HT levels simultaneously »	Evolution du 5HT dans le plasma lors du traitement par amitriptyline vs placebo	Wilcoxon's apired rank : Différence entre AUC (mol/L)		2.2 (1.1-4.9) x 10 ⁻¹⁹	3.6 (1.5-6.4) x 10 ⁻¹⁹ P < 0.001

PARTIE 2 : FIBROMYALGIE

PARTIE 2 : FIBROMYALGIE

I- INTRODUCTION :

1. Contexte :

La fibromyalgie touche 1 à 9% de la population mondiale, principalement des femmes (3.5% femmes pour 0.5% hommes) (1) (2) toute classe socioéconomique confondue. Bien que la pathologie soit suspectée depuis les années 70 (alors dénommée « fibrosite », « fibromyosite », « polyenthésopathie » ou « SPID (Syndrome polyalgique idiopathique diffus) ») puis reconnue en 1990 par le collège américain de rhumatologie et en 1992 par l'OMS, sa physiopathologie reste inconnue. L'hypothèse actuellement la plus probable est un dérèglement du système central de la douleur. Cette théorie est entre autres soutenue par la découverte d'une augmentation de la substance P dans le LCR des patients atteints de fibromyalgie (3), et par l'observation, à l'IRM cérébral, d'une activation des aires cérébrales de la douleur pour des stimuli d'intensité deux fois plus faibles que pour des sujets sains (4).

Bien qu'étant un motif fréquent de consultations en médecine générale (2 à 6 % des consultations), la prise en charge de ce syndrome est rendue difficile par la pluralité de ses symptômes : douleur chronique, troubles du sommeil, asthénie, etc..., ainsi que par l'absence de traitement étiologique et la mauvaise réponse aux traitements antalgiques. L'amitriptyline est le premier traitement ayant montré une efficacité et il est encore largement utilisé dans le monde. Pourtant dans une méta-analyse de 2015, *Amitriptyline for fibromyalgia in adults* R Andrew Moore et al ont montré une efficacité sur la douleur pour un quart des patients comparativement au placebo, avec un tiers d'entre eux éprouvant des effets indésirables non graves. Par ailleurs les auteurs signalent le faible niveau de validité scientifique des études attestant de l'efficacité de la molécule sur ce critère de jugement (5).

L'amitriptyline a l'AMM au Canada pour le traitement de la fibromyalgie (6), mais pas aux Etats Unis du fait de ses effets indésirables (AMM pour la duloxétine, la milnacipran et la prégabaline) (7), ni en Europe où aucun médicament n'a d'AMM pour cette pathologie. Ceci a d'ailleurs conduit en France à une prise en charge hétérogène des patients avec la prescription d'antalgiques (60%), d'antidépresseurs (40%), d'anxiolytiques-hypnotiques (30%) et d'antiépileptiques (20%) (8). En 2007, l'EULAR (Ligue Européenne Contre le Rhumatisme)

recommandait (grade A) l'usage d'antalgiques de pallier I et II (tramadol) et d'antidépresseurs (dont l'amitriptyline). Actuellement l'EULAR positionne la prise en charge médicamenteuse comme « faiblement approuvé » pour 5 molécules : pour le traitement de fond : amitriptyline, duloxétine, milnacipran et prégabaline, et pour le traitement de crises : (doliprane) tramadol. Elle ne retient plus aucune indication pour les autres médicaments (9).

2. Objectif :

Notre travail vise à vérifier le niveau de preuve de l'amitriptyline pour le traitement des patients adultes atteints de fibromyalgie, par une revue systématique de la littérature avec une méta-analyse des essais randomisés contre placebo, à travers 3 critères de jugements : la douleur, la qualité du sommeil et l'amélioration globale.

II- MATERIELS ET METHODE :

1. Protocole et enregistrement :

Le protocole de recherche de cette revue systématique de la littérature a été enregistré le 21 janvier 2022 sur la plateforme OSF sous le numéro d'immatriculation : 10.17605/OSF.IO/VA549

2. Critères d'éligibilité :

Les critères d'inclusion retenus pour cette revue de la littérature sont les suivants :

Les études doivent être des essais cliniques, randomisés, en double aveugle testant l'amitriptyline contre un placebo. Tous les schémas d'intervention sont éligibles aussi bien les études en parallèle qu'en cross-over. La population étudiée doit être une population adulte de plus de 18 ans.

Le critère de jugement principal de notre étude est la douleur. Les critères secondaires sont le sommeil et l'amélioration globale. Toutes les études évaluant ces paramètres en critère principal ou secondaire sont éligibles. Il n'y a pas de restriction quant à la posologie du traitement, sa durée ou l'année de publication de l'étude. Les articles doivent par ailleurs être rédigés en anglais ou en français.

3. Sources d'information :

La recherche des articles éligibles a été réalisée sur trois bases de données bibliographiques scientifiques, toutes consultées en ligne :

- ClinicalTrials.gov
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Medline (PubMed)

Dans un second temps, les études retenues par les auteurs ayant précédemment réalisé des revues systématiques de la littérature et/ou méta-analyses dans le même domaine de recherche ont été consultées, afin d'assurer une collecte optimale des articles éligibles. Une fois la recherche bibliographique terminée, une vérification des bibliographies des articles a été

réalisée afin de n'en omettre aucun. L'éligibilité des articles présents dans les méta-analyses et revues systématiques publiées dans pubmed et cochrane, ont été analysés de la même façon.

La recherche bibliographique s'est déroulée jusqu'à janvier 2020. Une veille bibliographique a été réalisée jusqu'en septembre 2021.

4. Recherche :

Le mot clé MeSH employé pour la recherche bibliographique est « AMITRIPTYLINE ». Le filtre « randomized controlled trial » a été ajouté pour la recherche d'articles dans les bases de données Cochrane CENTRAL et Medline (PubMed).

5. Sélection des études :

Le logiciel ZOTERO ® a été utilisé pour la sauvegarde et le tri des références bibliographiques. Les auteurs Houdna LAIB (H.L.) et Valérie GUICHARDANT (V.G.) ont réalisé indépendamment une recherche bibliographique exhaustive dans les 3 bases de données ci-dessus.

Dans un premier temps les auteurs ont utilisé "amitriptyline" comme mot clé de recherche dans PubMed et Cochrane avec le filtre "essai clinique randomisé". Elles ont sélectionné par la lecture des titres toutes les études concernant la fibromyalgie. Pour chaque article dont le titre a semblé pertinent, le résumé a été lu. Ceci a permis de déterminer si l'essai clinique remplit les critères d'inclusion. Si le résumé d'un article ne suffisait pas pour statuer, ce dernier a été lu en intégralité.

A la suite de ce travail, les deux auteurs ont confronté leur bibliographie. Les articles qu'elles ont toutes deux retenus ont été inclus ; les autres ont fait l'objet d'une réévaluation commune. En cas de désaccord persistant, un troisième intervenant indépendant a été sollicité pour trancher, le docteur Yves BARO (Y.B.) ou le professeur Rémy BOUSSAGEON (R.B.).

Une fois ces articles sélectionnés, les doublons d'une même étude ont été identifiés et éliminés.

Les textes intégraux des études sélectionnées ont été récupérés. Les auteurs ont alors analysé les bibliographies des articles sélectionnés, afin d'être exhaustives dans leur recherche bibliographique.

Le processus de sélection des articles est rapporté à l'aide d'un diagramme de flux selon les lignes directrices du protocole PRISMA 2020 (10).

6. Extraction des données :

Les données des études ont été extraites indépendamment par les deux auteurs (V.G. et H.L.) selon les caractéristiques suivantes :

- Méthode de l'étude : Type d'étude : parallèle ou cross-over ; durée de l'étude : durée totale pour les essais en parallèles, période de traitements et de wash-out pour les essais en cross-over ; type d'analyse réalisée (ITT, ITTm ou per protocole).
- Population étudiée : nombre de participants randomisé ; nombre de participants exclus et perdus de vue ; nombre de participants en fin d'intervention ; critères d'inclusion et d'exclusion ;
- Intervention de l'étude : nom des molécules testées dont amitriptyline et placebo ; posologies administrées pendant l'essai ; thérapies non médicamenteuses testées.
- Critères de jugement : critères principaux et secondaires
- Résultats de l'étude : significatifs ou non pour l'amitriptyline versus placebo
- Notes : utilisation ou non d'une méthode d'ajustement de la significativité en cas de critères de jugement multiples, conflits d'intérêts déclarés

7. Risque de biais inhérent à chacune des études

L'outil « Revised Cochrane risk-of-bias tool for randomized trials (RoB 2) » (11) a été appliqué indépendamment par les deux auteurs (V.G. et H.L.) pour chaque étude, afin d'évaluer sa qualité méthodologique. Les auteurs ont ensuite confronté leurs analyses de biais, et en cas de différence, le biais a été réanalysé. En cas de non-aboutissement à un consensus, les auteurs ont soumis l'analyse à un troisième intervenant indépendant pour statuer : le docteur Yves BARO (Y.B.) ou le professeur Rémy BOUSSAGEON (R.B.).

L'outil « RoB 2 » propose d'analyser la robustesse scientifique de chaque étude au travers de 5 catégories de biais, évalués chacune indépendamment en trois niveaux : risque faible, quelques inquiétudes et risque important :

- Biais résultants du processus de randomisation
- Biais résultants des écarts aux interventions prévues
- Biais résultants des données manquantes
- Biais résultants de la mesure du critère de jugement
- Biais résultants de la sélection des résultats rapportés

L'étude est considérée comme étant à faible risque de biais si chacune des 5 catégories est jugée à « faible risque de biais ». Elle est considérée comme présentant quelques inquiétudes si 4 catégories sur 5 sont à « faible risque de biais », mais que 1 catégorie sur 5 pose « quelques inquiétudes ». Elle est enfin considérée comme étant à haut risque de biais si au moins une catégorie de biais est jugée à « haut risque » ou si plusieurs catégories présentent « quelques inquiétudes ».

L'application de l'outil RoB2 sur les études incluses a suscité de nombreuses interrogations et parfois l'outil n'a pas apporté de réponse adéquate à ces questions. Les deux auteurs ont alors instauré avec accord du Pr. Boussageon des règles générales permettant de statuer plus facilement.

Premièrement, pour les études présentant des critères de jugement multiples, des méthodes d'ajustement permettant le contrôle du risque d'erreur de type I seront recherchées. En l'absence d'ajustement du risque alpha, l'analyse sera jugée inadéquate même si une différence statistique est retrouvée.

Deuxièmement, pour les différentes catégories de biais :

Concernant le “Biais résultats du processus de randomisation” :

- Il a été statué que lorsqu'une étude est de grande envergure, par exemple réalisée dans plusieurs centres universitaires ou que l'équipe d'investigation a réalisé plusieurs publications précisant dans l'une d'elle une méthode de randomisation valide ; bien qu'il n'existe pas de détails concernant la séquence d'allocation ni de sa dissimulation

dans l'étude analysée ; la réponse à l'item sera "probablement oui" plutôt que "absence d'informations".

- Lorsqu'il s'agit d'études en cross-over, il a été décidé que lorsque les auteurs de l'essai ne présentent pas les caractéristiques de base de la population initiale la réponse à l'item sera "absence d'information" plutôt que "probablement non".

Concernant le "Biais résultants des écarts aux interventions prévues" :

- Lorsque l'étude exclut uniquement les perdus de vue, celle-ci sera considérée comme utilisant une analyse en intention de traiter modifiée (ITTm). Mais si une étude exclut des participants pour tout autre raison, alors celle-ci sera considérée comme utilisant une analyse en per protocole.
- Le double aveugle est considéré comme levé lorsqu'il existe plus de 50% d'effets indésirables spécifiques à l'amitriptyline définis par somnolence ou sécheresse buccale dans le groupe traité comparé au groupe placebo (12-13)
- Lorsque dans les essais cliniques les effets indésirables spécifiques à l'amitriptyline ne sont pas rapportés, les auteurs (H.L. et V.G.) ont considéré qu'il était difficile de statuer sur le maintien du double aveugle et dans ce cas la réponse à l'item sera "absence d'information".
- Une analyse en ITTm est jugée comme inadéquate s'il existe plus de 5% de données manquantes.
- Lorsque l'analyse employée est jugée inadéquate, il a été décidé d'estimer systématiquement qu'il existe un risque d'une répercussion sur le résultat.

Concernant le " Biais résultants des données manquantes " :

- Le seuil des données manquantes est fixé à 5%.
- Lorsque les causes de données manquantes sont mal précisées, l'estimation de leur possible impact sur la vraie valeur sera jugée " absence d'information".

Concernant le " Biais résultants de la mesure du critère de jugement " :

- Lorsque les critères de jugement sont considérés comme étant subjectifs, et que le double aveugle est considéré comme levé alors l'estimation d'une possible influence de l'évaluation par la connaissance du groupe d'intervention est jugée "probablement oui"

- Il a été décidé de ne pas se positionner concernant l'influence de la connaissance de l'intervention reçue sur l'évaluation des résultats avec un jugement porté "absence d'informations".

Concernant le "Biais dans la sélection des résultats rapportés" :

- Lorsqu'il n'est pas fait mention que le protocole de l'étude a été enregistré préalablement, il a été statué "absence d'informations".
- Lorsque les auteurs de l'essai annoncent avoir trouvé une différence significative sans donner le résultat il a été décidé qu'il y avait un risque de sélection de résultat avec un jugement porté : "probablement oui".

Concernant le risque d'inflation du risque alpha dans le cas de multiplicité des tests :

- Lorsque les auteurs rapportent avoir réalisé un ajustement lors d'une précédente étude, il a été statué que pour l'étude en question il est considéré que l'ajustement a également été réalisé.

Etant donné que nous avons statué :

- La levée du double aveugle en cas d'absence d'information sur les effets indésirables spécifiques de l'amitriptyline
- Un seuil de données manquante et d'acceptabilité de perdus de vue analyse en ITTm à 5%
- Une réponse "absence d'information" si les auteurs ne déclarent pas avoir enregistré leur protocole au préalable

Et que ces décisions impactent fortement l'évaluation du risque de biais de chaque catégorie, il a été décidé avec Pr. Boussageon de réaliser une étude de sensibilité méthodologique lorsque cela est pertinent. Ainsi, pour chaque article scientifique analysé, nous regarderons si son niveau global de preuve est impacté par UNE SEULE de ses trois décisions.

Si c'est le cas, l'article sera réanalysé avec :

- Pour la non mention des effets indésirables une estimation d'absence d'impact sur le double aveugle
- Pour les données manquantes et le seuil d'acceptabilité de perdu de vue d'ITTm à 10 et 15%

- Une réponse “ probablement oui “ si les auteurs ne fournissent aucune information sur l’enregistrement du protocole

8. Quantification des résultats :

Les auteurs ont retenu la douleur comme critère de jugement principal, et le sommeil et l’amélioration globale du patient comme critères de jugement secondaires. La mesure de ces résultats est dans la majorité des études réalisée à l’aide d’une échelle visuelle analogique. Les résultats sont de ce fait présentés sous forme de moyennes $\pm$ l’écart-types (SD).

9. Synthèse des résultats :

L’ensemble des études incluses sera analysé, afin de constituer une synthèse qualitative. Si suffisamment de résultats sont éligibles, sous forme de moyenne avec écart type ; une synthèse quantitative sera également réalisée, à l’aide du logiciel Review Manager version 5.4 (RevMan[®]) selon les préconisations de la méthodologie REB (14).

L’analyse quantitative sera conditionnée par la gestion des données manquantes. Pour ce faire, les auteurs des ECR et des méta-analyses ont été sollicités par mail, afin qu’ils nous transmettent leurs recueils de données. Des relances hebdomadaires seront réalisées sur la période d’un mois. Les articles pour lesquels la moyenne $\pm$ l’écart type d’aucun des trois critères de jugement ne pourront être obtenu seront exclus de l’analyse quantitative.

Selon la diversité des échelles de mesures de nos critères de jugement dans les articles, des différences de moyennes (si besoin standardisées) seront calculées à partir des résultats ainsi que leurs intervalles de confiance à 95%. Le risque alpha sera fixé à 0,05. L’hétérogénéité des résultats des essais sera par la suite évaluée grâce à un test statistique de Chi2. Un modèle d’effet fixe sera utilisé préférentiellement. En cas d’hétérogénéité forte ($I^2 > 50 \%$) un modèle d’effet aléatoire sera utilisé.

Dans la mesure du possible, deux méta-analyses seront réalisées pour pouvoir répondre au protocole REB (14) et parvenir à une conclusion.

Une méta-analyse principale, qui ne prendra en considération que les résultats d'essais cliniques randomisés (ECR) de haute qualité. Le résultat de cette méta-analyse sera considéré comme positif si les résultats sont statistiquement significatifs ($p < 0,05$). Le nombre d'ECR concluants selon l'analyse du RoB 2, l'hétérogénéité et le biais de publication entre les ECR seront évalués. Une autre méta-analyse dite "de sensibilité" sera réalisée incluant tous les ECR pour lesquels les résultats sont analysables, quelle que soit leur qualité, permettant de vérifier la robustesse des résultats de la méta-analyse de base.

10. Risque de biais transversal aux études et évaluation du niveau de preuve :

Le niveau de preuve des résultats de la méta-analyse sera évalué selon l'approche GRADE (Grading of recommendations assessment, development and evaluation)(15).

Le niveau de preuve des résultats issus d'une méta-analyse d'ECR est en principe le plus élevé, mais il peut être diminué dans plusieurs situations : en cas de risque de biais dans les ECR, d'imprécision (intervalle de confiance trop large), de forte hétérogénéité (résultats contradictoires), de preuve "indirecte" (population, intervention, critère ou comparateur non totalement adapté) ou de présence d'un biais de publication. Ainsi selon les caractéristiques remplies le niveau de preuve pourra être défini comme "élevé", "modéré" ou "faible".

Dans le cadre de la méthodologie REB une conclusion finale sera proposée en fonction de la configuration de l'étude et de plusieurs critères :

- Si une méta-analyse principale trouve des résultats positifs : trois conclusions sont possibles selon 4 critères (Nombre d'ECR concluants, Niveau d'hétérogénéité, Biais de publication et Résultats de la méta-analyse de sensibilité).
- Si une méta-analyse principale n'est pas possible mais qu'une analyse de sensibilité est possible : trois conclusions sont possibles en fonction de 3 critères (nombre d'ECR concluants, Biais de publication, Résultats de la méta-analyse de sensibilité).
- Si ni l'analyse principale ni l'analyse de sensibilité ne sont possibles : trois conclusions sont possibles en fonction de 2 critères (nombre d'ECR concluants et biais de publication).

Les trois conclusions possibles sont :

- Preuve solide
- Preuve à confirmer
- Absence de preuve

III- SELECTION DES ETUDES :

La recherche avec le seul mot clé « amitriptyline » et l'usage dans PubMed du filtre « essai clinique randomisé » a mis en évidence 3418 articles dont 2541 dans Cochrane, 760 dans Medline et 117 dans Clinical Trials. Les articles dont les titres et abstracts ne remplissaient pas les critères d'inclusion récapitulés ci-dessous ont été exclus :

- Essai clinique randomisé
- En double aveugle
- Comparant l'amitriptyline à un placebo
- Rédigés en français ou en anglais
- Dont le texte intégral est disponible
- Ayant comme objectif principal ou secondaire l'évaluation de la douleur / de la fatigue / de la qualité du sommeil / de l'amélioration globale du patient

107 articles correspondaient par leur titre et abstracts aux critères d'inclusion : 61 articles dans Cochrane, 32 dans Medline et 14 dans Clinical Trials.

Parmi ces articles, il a été impossible de récupérer le texte intégral pour 3 d'entre eux, 1 était rédigé dans une langue autre que l'anglais ou le français, 3 se sont révélés être hors sujet et 50 étaient des doublons.

Nous avons pu inclure 13 articles pour notre analyse :

A - S. Carette 1986: Evaluation of amitriptyline in primary fibrositis. A double-blind, placebo-controlled study (16).

B - Don. Goldenberg 1986: A randomized, controlled trial of amitriptyline and naproxen in the treatment of patients with fibromyalgia (17).

C - Ra. Scudds 1989: Improvements in pain responsiveness in patients with fibrositis after successful treatment with amitriptyline (18).

D - R. Jaeschke 1991: Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials (19).

E - S. Carette 1994: Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. A randomized, double-blind clinical trial (20).

F - C. Kempenaers 1994: Effect of an antidiencephalon immune serum on pain and sleep in primary fibromyalgia (21).

G - S. Carette 1995: Sleep electroencephalography and the clinical response to amitriptyline in patients with fibromyalgia (22).

H - Don. Goldenberg 1996: A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia (23).

I - F. Ginsberg 1996: A randomized placebo-controlled trial of sustained-release amitriptyline in primary fibromyalgia (24).

J - P. Hanonnen 1998: A randomized double-blind placebo controlled study of moclobemide and amitriptyline in fibromyalgia (25).

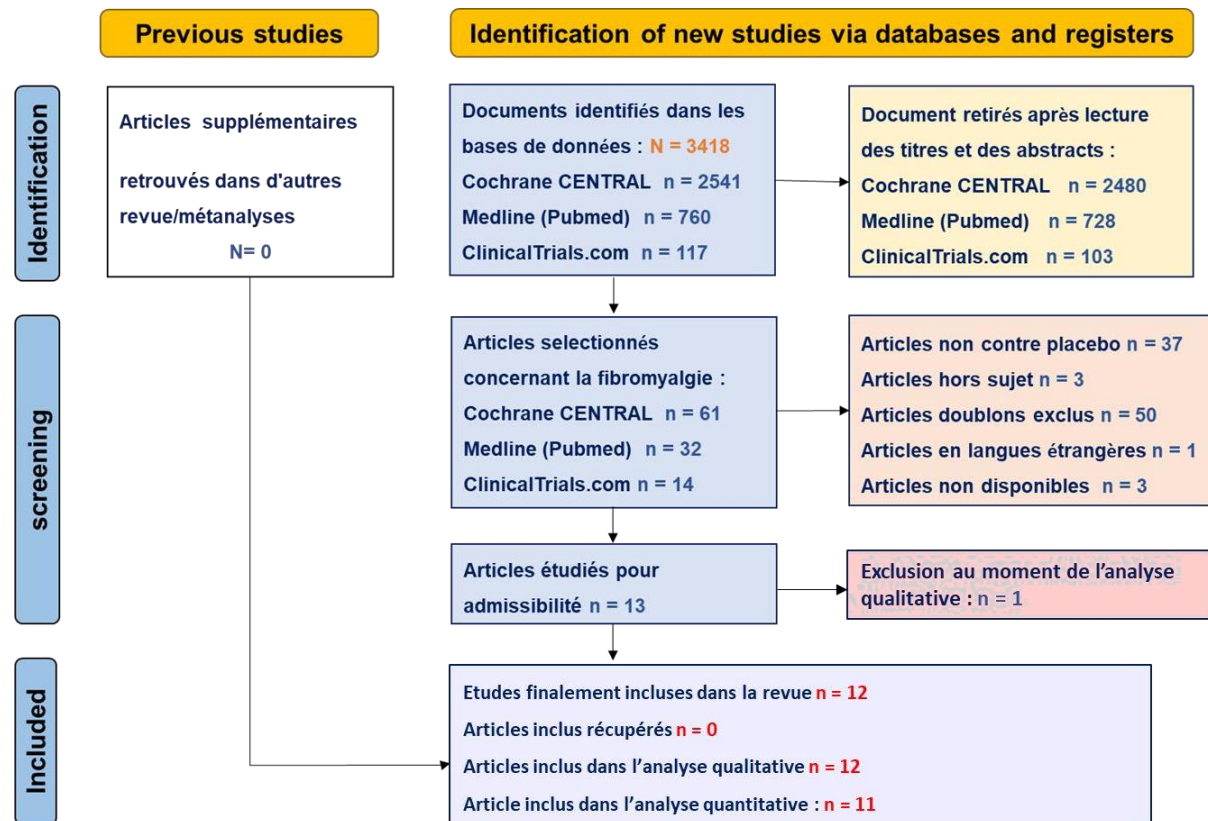
K - R.E Heymann 2001: A double-blind, randomized, controlled study of amitriptyline, nortriptyline and placebo in patients with fibromyalgia. An analysis of outcome measures (26).

L - A Fors 2002: The effect of guided imagery and amitriptyline on daily fibromyalgia pain: a prospective, randomized, controlled trial (27).

M - A. Braz 2012: Effects of Panax ginseng extract in patients with fibromyalgia: a 12-week, randomized, double-blind, placebo-controlled trial (28).

L'étude R. Jaeschke 1991: *Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials* a été secondairement exclue, au moment de l'analyse qualitative. En effet, sa méthodologie est particulière : type d'essai en cross-over en « single case » (cas singulier) ou « N-of one » (N de 1). De plus les auteurs n'ont pas rapporté pour nos critères de jugement les résultats en faveur du traitement par placebo (présentation uniquement du nombre de patients significativement améliorés sous amitriptyline). Les résultats de l'étude sont donc incomplets par rapport à notre travail de recherche et ont été jugés inexploitable pour l'analyse qualitative et pour l'analyse quantitative. Les informations de l'étude : caractéristiques, risque de biais, ajustement du risque de biais selon les interprétations de 3 items et résumé, sont rapportées à titre indicatif, car l'essai clinique remplit les critères d'inclusion et d'exclusion de notre projet de recherche. Les données de l'article n'ont été prises en compte dans aucune analyse : ni dans l'analyse de sensibilité du risque de biais, ni dans les analyses qualitatives ou quantitatives des résultats.

Ainsi, 12 études ont été incluses et ont permis une synthèse qualitative. 11 d'entre elles ont été incluses dans l'analyse quantitative : Carette 1986, Scudds 1989, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Heymann 2001, Fors 2002, Braz 2012 (voir Figure 1).



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Figure 1 : flow chart selon le modèle PRISMA (10) pour la fibromyalgie

IV- RESULTATS

1. Caractéristiques des études sélectionnées

12 études ont été incluses pour l'analyse qualitative. Elles ont été publiées entre 1986 et 2012 et regroupent au total 877 patients randomisés, pour 708 patients analysés, issus de 6 pays : Canada, USA, Belgique, Finlande, Brésil, Norvège. Les caractéristiques de l'étude Jaeschke 1991 ont aussi été rapportées à titre indicatif.

a) Schéma d'étude :

9 Etudes sont des essais en groupes parallèles (Carette 1986, Don Goldenberg 1986, Carette 1994, Kempenaers 1994, Ginsberg 1996, Hanonnen 1998, Heymann 2001, Fors 2002, Braz 2012) avec un nombre de participants variant de 36 (Kempenaers 1994) à 208 (Carette 1994).

3 Etudes sont des essais en cross-over (Scudds 1989, Carette 1995, Don Goldenberg 1996) avec un nombre de participants variant de 22 (Carette 1995) à 39 (Scudds).

3 études sont multicentriques : Carette 1986, Carette 1994 et Hanonnen 1998. Les 9 autres études sont monocentriques.

Toutes les études sélectionnées comparent l'amitriptyline à un placebo inactif. Cependant 2 études ne comparent pas les résultats finaux obtenus respectivement sous amitriptyline et sous placebo, mais mesurent l'amélioration intragroupe uniquement, sans recours à un test statistique de comparaison : Kempenaers 1994, et Ginsberg 1996. L'étude de Don Goldenberg 1996 suggère avoir eu recours à une comparaison des résultats finaux de l'amitriptyline contre placebo, mais les résultats présentés évoquent une comparaison des améliorations intragroupes (voir le chapitre « résumé »).

7 essais cliniques étudient aussi l'efficacité de l'amitriptyline par rapport à une molécule active : le naproxen pour Don Goldenberg 1986, la cyclobenzaprine pour Carette 1994, le SER282 pour Kempenaers 1994, la fluoxétine pour Don Goldenberg 1996, le moclobemide pour Hanonnen 1998, la nortriptyline pour Heymann 2001 et le p. ginseng pour Braz 2012.

A titre indicatif, l'étude Jaeschke 1991 utilise un schéma cross-over particulier : « single case ou N-of-1 » et compte 23 essais pour 22 participants.

b) Durée des études

Les essais cliniques ont duré entre 4 et 32 semaines. L'efficacité de l'amitriptyline a été testée sur des périodes de traitement allant de 4 semaines à 6 mois.

Pour les essais en parallèle, la durée minimale a été de 4 semaines (Fors 2002) et la durée maximale de 6 mois (Carette 1994). On notera qu'un essai a duré 6 semaines (Don Goldenberg 1986), 3 essais ont duré 8 semaines (Kempenaers 1994, Ginsberg 1996, Heymann 2001), un essai a duré 9 semaines (Carette 1986), 2 essais ont duré 12 semaines (Hanonnen 1998 et Braz 2012).

Pour les études en cross-over, la durée des essais cliniques est respectivement de 10 semaines (Scudds 1986), 16 semaines (Carette 1995), et 32 semaines (Don Goldenberg 1996). La durée de chaque période de traitement a été respectivement de 4 semaines (Scudds 1989), 6 semaines (Don Goldenberg 1996) et 8 semaines (Carette 1995). A noter que les deux essais qui ont les plus petites périodes de traitement ont été réalisés chacun avec une période de wash out de 2 semaines.

1 étude a intégré à son schéma une période de prétraitement : Hanonnen 1998 : essai en simple aveugle visant à exclure les patients répondant au placebo.

A titre indicatif, dans l'étude Jaeschke 1991, la durée de suivi de chaque patient a été personnalisée. L'essai n'a pas intégré de période de wash out. Chaque période de traitement a duré 2 semaines. Le protocole d'étude a intégré une période de prétraitement : essai ouvert recherchant la dose individuelle d'amitriptyline.

c) Caractéristiques des populations étudiées

Les 12 études ont inclus uniquement des patients majeurs et atteints de fibromyalgie, selon les critères ci-dessous :

- Pour 8 études : critères de American College of Rheumatologie de 1990 : Carette 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Heymann 2001, Fors 2002 et Braz 2012.

- Pour 2 études : critères de Yunus et Al : Kempnaers 1994 et Don Goldenberg 1986
- Pour 2 études : critères de Smythe et Moldofsky : Scudds et Carette 1986.

Les patients atteints de pathologies cardiaques ou rénales, ainsi que ceux ayant des antécédents de rétention d'urine ou de glaucome et ceux présentant un état de grossesse ou étant en période d'allaitement maternel ont systématiquement été exclus. Ceux souffrant de pathologies pulmonaires, neurologiques, infectieuses, rhumatismales, endocrinologiques, digestives, ont le plus souvent aussi été exclus. 3 études ont également exclu les patients atteints de pathologies psychiatriques sévères : Hanonnen 1998 et Kempnaers 1994, ou de dépression sévère Don Goldenberg 1996. 5 études n'ont accepté aucun patient prenant un autre traitement que le paracétamol pendant l'essai clinique : Carette 1986, Carette 1994, Carette 1995, Don Goldenberg 1986 et Scudds 1989. 4 études ont exclu les patients prenant des « traitements actifs sur les SNC », des AINS ou des antalgiques autres que le doliprane : Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Heymann 2001. 1 étude a pour seul critère d'exclusion rapporté : sexe masculin : Fors 2002.

La moyenne d'âge des patients ayant participé aux essais cliniques a varié de 38 ans (Kempnaers 1994) à 51 ans (Heymann 2001).

La population étudiée est principalement féminine : entre 824 et 832 femmes pour 877 patients (deux études ne fournissent que le nombre de femmes en analyse finale). 5 études n'ont inclus que des femmes : Kempnaers 1994, Hanonnen 1998, Heymann 200, Fors 2002 et Braz 2012 soit 397 patientes. Les pourcentages de femmes dans les 7 autres études ont varié de 83% à 96%.

La durée moyenne des symptômes de fibromyalgie des patients a été mesurée pour 9 études : Carette 1986, Carette 1994, Kempnaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Fors 2002 et Braz 2012. Elle a varié de 38 mois (Ginsberg 1996) à 187 mois (Fors 2002) avec une durée moyenne de 91 mois.

A titre indicatif, l'étude Jaeschke 1991, ne rapporte ni les critères d'inclusion ni d'exclusion de ses patients, hormis qu'ils doivent être atteints de fibromyalgie (critères diagnostics non précisés). La population est composée de 21 femmes pour 22 patients (bien que l'étude rapporte 23 essais en N-of-1, une femme ayant réalisé 2 essais), d'âge moyen 55 ans. La durée moyenne des symptômes de fibromyalgie n'est pas rapportée par les auteurs.

d) Analyse des résultats

Dans ces 12 études, le nombre de patients perdus de vue ou exclus est de 169 pour 877 patients randomisés soit 19 % de la population initiale.

Pour 9 essais cliniques, les auteurs ont eu recours à une analyse en intention de traiter modifiée, c'est-à-dire avec exclusion de l'analyse finale uniquement des patients perdus de vue. 1 étude a un pourcentage de perdus de vue inférieur ou égal à 5% : Fors 2002(5%). 5 études ont un pourcentage de perdus de vue inférieur ou égal à 10% : Don Goldenberg 1986 (7%), Scudds 1989 (8%), Carette 1994 (11.5%), Carette 1995 (9%), Ginsberg 1996 (10%) et Heymann 2001 (10%). Les 3 autres études ont un pourcentage de perdus de vue supérieur à 10% : Carette 1994 (11.5%), Braz 2012 (27%) et Don Goldenberg 1996 (39%).

Pour 1 essai clinique, l'analyse est en intention de traiter (tout patient ayant débuté le traitement assigné est analysé) : Hanonnen 1998. Les auteurs utilisent une méthode non valide pour la gestion de leurs perdus de vue (dernière observation réalisée reportée en résultat final) (29% de la population randomisée)

Pour 2 essais cliniques, les auteurs ont eu recours à une analyse en per protocole, c'est-à-dire avec exclusion de l'analyse finale des patients non observants et des patients perdus de vue. Il s'agit de Carette 1986 et Kempnaers 1994 avec respectivement 16% et 36% de perdus de vue.

A noter que l'étude de Carette 1994 a aussi été analysée en per protocole, mais que le résultat n'a pas été rapporté dans l'article par les auteurs car il a été jugé similaire à celui obtenu par analyse en ITTm.

A noter aussi que 2 études ne comparent pas les résultats finaux des patients traités par amitriptyline et par placebo, mais les améliorations intragroupes respectives : Kempnaers 1994 et Ginsberg 1996, ce qui semble aussi être le cas pour l'étude Don Goldenberg 1996 (voir chapitre « résumé »).

Enfin, à titre indicatif l'étude Jaeschke 1991 a été analysée en intention de traiter, avec une absence de perdu de vue.

e) Critères de jugement

L'évaluation globale des symptômes de la fibromyalgie par le patient est mesurée dans 11 études dont 2 en critère de jugement principal : Carette 1986 et Don Goldenberg 1996. Il a été décidé de mesurer ce critère de jugement selon 2 approches différentes : le ressenti global du patient et les répercussions de la pathologie sur sa qualité de vie mesurées par le Fibromyalgie Impact Questionnaire (FIQ) : Don Goldenberg 1996, Heymann 2001, Braz 2012. Le ressenti global est principalement mesuré par EVA : Don Goldenberg 1986, Carette 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, mais aussi par le VSGI (Heymann 2001) et un score d'amélioration de 1 à 5 avec résultat rapporté en nombre ou pourcentage de patients améliorés (Carette 1986 et Scudds 1989). Une étude ne précise pas la méthode de mesure utilisée : Kempenaers 1994.

L'évaluation globale des symptômes de la fibromyalgie par le médecin est mesurée dans 8 études dont 1 en critère de jugement principal : Hanonnen 1998. Elle est principalement mesurée par EVA : Don Goldenberg 1986, Carette 1994, Carette 1995, Don Goldenberg 1996 et Ginsberg 1996. Elle est aussi mesurée de 2 façons différentes dans l'étude Hanonnen 1998 : par la *CIC Impression Clinique de Changement* (critère de jugement principal) et par la *CIS Physician's Clinical impression of Severity* (critère de jugement secondaire). L'étude Carette 1986 mesure ce critère de jugement par un score d'amélioration de 1 à 5 avec résultat rapporté en nombre ou en pourcentage de patients améliorés. Une étude ne précise pas la méthode de mesure utilisée : Kempenaers 1994.

La douleur est mesurée indépendamment dans 11 études, dont une en critère de jugement principal : Scudds 1989. Dans les autres études, c'est un critère de jugement secondaire. Elle est principalement mesurée par EVA (Carette 1986, Don Goldenberg 1986, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Fors 2002 et Braz 2012), mais aussi par le score de Mc Gill (Scudds 1989 et Carette 1994).

Le sommeil est mesuré indépendamment dans 9 études, principalement par EVA (Don Goldenberg 1986, Kempenaers 1994, Carette 1994, Carette 1995, Don Goldenberg 1996,

Ginsberg 1996, Hanonnen 1998 et Braz 2012), mais aussi par pourcentage de patients améliorés sur une échelle : « pire », « inchangée », « mieux » (Carette 1986).

La fatigue est mesurée indépendamment dans 8 études, principalement par EVA (Don Goldenberg 1986, Carette 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998 et Braz 2012), mais aussi par un score de 0 à 3 (Kempenaers 1994).

Les autres critères de jugement mesurés par les auteurs sont :

- Evaluation des incapacités et du retentissement de la fibromyalgie :
 - *Sickness Impact Profile* SIP (Profile d'Impact de Maladie) : Carette 1994
 - *Health Assessment Questionnaire* HAQ (Questionnaire d'Evaluation de la santé) : Carette 1994
 - *Sheehan's functional scales* (Echelle Fonctionnelle de Sheehan) : Hannonen 1998
 - *Nottingham Health Profile dimension* (Profile de Santé selon les Dimension de Nottingham): Hannonen 1998
- Evaluation anxiété et dépression par
 - *Arthritis Impact Measurement Scales* AIMS (Echelle de mesure de l'Impact de l'Arthrite) : Carette 1994
 - Hamilton dépression et anxiété : Kempenaers 1994
 - *Beck depression Inventory* (Inventaire de dépression de Beck): Goldenberg 1996, Fors 2002
 - *STAI-T State-Trait Anxiety Inventory Trait* (Inventaire des Traits de l'Anxiété) : Fors 2002
 - *ATQ – 30 : Automatic negative Thoughts Questionnaire* (Questionnaire des Pensées Automatiques Négatives) : Fors 2002
 - EVA anxiété : Braz 2012
- Critère de jugement composite :
 - Sommeil, fatigue matinale, fatigue, crampes, dysménorrhée, constipation, céphalées, vertiges, gonflement des doigts : Kempenaers 1994 (critère de jugement secondaire)

- Évaluation du TMS, du seuil de douleur, de la tolérance de la douleur, et de l'intensité de la douleur (via le score de Mc Gill pour cette dernière) : Scudds 1989 (critère de jugement principal)
- Amélioration d'au moins 4 symptômes sur 6 avec majoration d'1 kg au moins du total myalgique score ou 50% d'amélioration sur les critères : douleur, sommeil, fatigue, évaluation globale par le patient et ou le médecin, mesurées par EVA : Carette 1994 et Carette 1995 (critère de jugement principal)
- Amélioration d'au moins 3 critères sur 4 : 50% ou plus sur l'évaluation globale des symptômes par le patient ou le médecin examinateur, 50% sur la douleur, 25% sur le TMS : Ginsberg 1996 (critère de jugement principal)
- Polysomnographie : Kempnaers 1994, Carette 1995
- Raideur matinale : Carette 1986, Carette 1994, Ginsberg 1995
- Nombre de points sensibles : Carette 1986, Kempnaers 1994, Ginsberg 1996, Hanonnen 1998, Heymann 2001, Braz 2012
- Score total de douleur musculaire (Total Myalgic Score : TMS) : Carette 1986, Scudds 1989, Don Goldenberg 1986, Carette 1994, Kempnaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996,
- Douleurs diffuses : Kempnaers 1994
- Tolérance douleur : Scudds 1989
- Seuil de douleur : Scudds 1989
- Score de sévérité de la douleur localisée au niveau cervical, scapulaire, lombaire, pelvis, membres supérieurs et inférieurs : Kempnaers 1994
- Sensation au réveil : Don Goldenberg 1986, Carette 1994, Don Goldenberg 1996, Ginsberg 1996

A titre indicatif, l'étude Jaeschke 1991 évalue 2 critères de jugement principaux : d'une part un score composite qui mesure entre autres la douleur, les troubles du sommeil, et la fatigue et d'autre part le nombre de points sensibles. Les critères de jugement secondaires sont : les critères de jugement du score composite étudiés de façon indépendante, le niveau de confiance du médecin dans sa Conduite A Tenir (CAT) avant et après essai et l'impact de l'essai sur cette dernière.

2. Risque de biais

Après évaluation des domaines de jugement du RoB 2 (biais résultant du processus de randomisation, de l'effet de période ou carry-over pour les essais en cross over, des écarts aux interventions, des données manquantes, de la mesure des résultats, et de la sélection des résultats rapportés) et du biais résultant d'une inflation du risque alpha, les 12 études incluses dans notre méta-analyse sont jugées à risque de biais « élevé ». A titre indicatif, l'étude Jaeschke 1991 a aussi un risque de biais global estimé élevé.

Le tableau ci-dessous récapitule pour chaque étude, son niveau de risque de biais par domaine (tableau 1) :

Nom l'auteur principal et année de publication	Biais résultant du processus de randomisation	Biais dû à l'effet de période et à l'effet carry over	Biais résultant du aux écarts aux interventions prévues	Biais résultant des données manquantes	Biais résultant de la mesure du critère de jugement	Biais résultant de la sélection des résultats rapportés	Biais résultant de l'inflation du risque alpha	Evaluation globale
Carette S. 1986								Risque élevé
Don Goldenberg 1986								Risque élevé
RA. Scudds 1989								Risque élevé
R. Jaeschke 1991								Risque élevé
S. Carette 1994								Risque élevé
C. Kempenaers 1994								Risque élevé
S. Carette 1995								Risque élevé
Don Goldenberg 1996								Risque élevé
F. Ginsberg 1996								Risque élevé
P. Hanonnen 1998								Risque élevé
R.E Heymann 2001								Risque élevé
E.A Fors 2002								Risque élevé
A. Braz 2012								Risque élevé

Tableau 1 : Représentation du niveau de risque de biais par domaine sur les 12 études incluses, ainsi que Jaeschke 1991

L'évaluation du pourcentage d'études présentant un risque de biais jugé « faible » ou « présentant quelques inquiétudes » ou « élevé » par domaine de biais a été réalisée pour les 12 études incluses dans notre méta-analyse (voir figure 2)

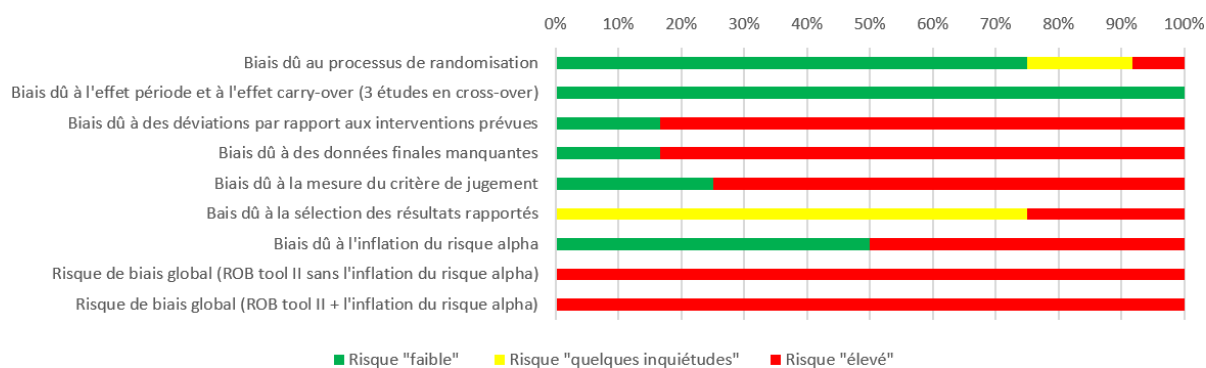


Figure 2 : Evaluation par domaine de biais du pourcentage d'études présentant un risque de biais estimé « faible » ou présentant « quelques inquiétudes » ou « élevé » parmi les 12 études de notre méta-analyse.

Le graphique (figure 2) montre que les 2 domaines de biais estimés à risque « élevé » les plus fréquents sont : les biais liés aux déviations par rapport aux interventions prévues et les biais liés aux données manquantes (83.33% des études). Ensuite viennent les biais liés à la mesure du critère de jugement (75% des études), puis à la sélection des résultats rapportés (25% des études) et enfin au processus de randomisation (8.33% des études). A noter que toutes les études sont à risque de biais « faible » pour l'effet carry-over et pour l'effet période, mais cela ne concerne que les 3 essais cliniques réalisés en cross-over. L'outil RoB 2 de Cochrane ne prend pas en compte le biais résultant de l'inflation du risque alpha contrairement à la méthodologie REB. Il est estimé élevé pour 50% des études.

La description de l'évaluation des domaines de biais pour chacune des 12 études sélectionnées, ainsi que pour Jaeschke 1991, est présentée ci-dessous.

a) Les biais dans le processus de randomisation :

Pour 9 études, le risque biais résultant du processus de randomisation est estimé « faible » : Braz 2012, Fors 2002, Heymann 2001, Hanonnen 1998, Don Goldenberg 1986 et 1996, Carette 1995, 1994 et 1986. La méthode de maintien de la dissimulation de la séquence d'allocation jusqu'à inscription et répartition des participants aux interventions est décrite dans 6 études : « méthode centralisée d'attribution des interventions aux participants » pour Braz 2012, Heymann 2001, Hanonnen 1998, Don Goldenberg 1996 et « contenant scellés jusqu'au début

de l'intervention » Fors 2002 et Don Goldenberg 1986. Le processus d'allocation aléatoire est décrit dans 7 études : « *table random number* » pour Carette 1994, Carette 1995, Don Goldenberg 1996, Heymann 2001, « *sequential numbered envelope* » pour Hanonnen 1998, « choix en aveugle par le patient d'un contenant de traitement » pour Fors 2002 et « attribution à chaque patient inclus dans l'étude d'une enveloppe contenant aléatoirement l'un des traitements de l'essai » Don Goldenberg 1986. Les caractéristiques de base des patients entre les groupes d'intervention au début de l'étude ont été comparées sans mettre en évidence de différence qui ferait évoquer un problème au niveau du processus de randomisation dans 3 études : Heymann 2001, Hanonnen 1998, Don Goldenberg 1986. 2 études n'ont pas réalisé de comparaison entre leurs groupes d'intervention car il s'agit d'essais en cross-over : Carette 1995 et Don Goldenberg 1996. 3 autres études ont utilisé une méthode non valide de comparaison de l'homogénéité des groupes d'intervention : 2 études ont réalisé une analyse des caractéristiques des patients ayant terminé l'étude : Braz 2012 et Carette 1986 et 1 étude n'a pas tenu compte d'un facteur pronostique dans la comparaison des caractéristiques de base de ses groupes d'intervention : Fors 2002. L'étude de Carette 1986 est un grand essai clinique, multicentrique, mené dans une unité d'essais cliniques expérimentée et nous avons donc considéré que le processus de randomisation : allocation aléatoire et maintien de la dissimulation de séquence, a été réalisé de façon adéquate, bien que les auteurs ne précisent pas leur méthode et que l'analyse de comparabilité des groupes d'intervention ait été réalisée sur la population finale de l'essai et non sur celle présente lors de la randomisation. De la même façon, bien que les essais de Carette 1994 et Carette 1995 ne précisent pas leur méthodologie de maintien du secret de l'allocation, il est probable, devant l'expérience des auteurs, que la randomisation ait été réalisée de façon adéquate.

Pour 2 études, les articles mentionnent qu'une randomisation a été réalisée, sans préciser ni la méthode utilisée, ni la comparaison des caractéristiques des patients dans les groupes d'intervention ce qui conduit à estimer un risque de biais intermédiaire : Scudds 1989, Ginsberg 1996.

Pour 1 étude, le risque de biais lié à la randomisation est jugé élevé : Kempnaers 1994. La méthode de randomisation n'est pas rapportée dans l'article et les auteurs rapportent des différences de base entre les groupes d'intervention suggérant un biais dans le processus.

A titre indicatif, l'étude Jaeschke 1991 mentionne avoir réalisé une randomisation, sans décrire la méthode utilisée.

b) Les biais dus à l'effet de période et à l'effet carry-over

3 études sont des essais en cross-over : Scudds 1989, Carette 1995, Don Goldenberg 1996. L'évaluation du risque de biais dans ce domaine a été jugé faible pour chacune. Dans les 3 études, le ratio de participants dans chaque bras de chaque période de traitement est de 1 pour 1. L'amitriptyline a une demi-vie d'élimination plasmatique comprise entre 22 et 40 heures selon les individus et ne semble pas avoir d'effet carry-over.

A titre indicatif, l'essai clinique Jaeschke 1991 est un essai en cross-over particulier qui alterne sans période de wash out les traitements par amitriptyline et par placebo sur des périodes de 2 semaines. Les auteurs justifient cela par le fait qu'ils estiment que l'amitriptyline n'a pas d'effet carry-over.

c) Les biais dus aux écarts aux interventions prévues

Pour 10 études : Carette 1986, Don Goldenberg 1986, Scudds 1989, Carette 1994, Kempnaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Heymann 2001 et Braz 2012, le risque de biais est jugé élevé.

2 études sont à faible risque de biais pour ce domaine : Hanonnen 1998 et Fors 2002. En effet, pour ces 2 essais cliniques, il n'est pas possible de statuer sur le maintien ou non du double aveugle, étant donné que les effets indésirables n'ont pas été mesurés ou leurs pourcentages non rapportés. Il n'a pas été constaté d'écart aux interventions prévues par leurs protocoles respectifs. L'analyse utilisée est adaptée : intention de traiter pour Hanonnen 1998, et intention de traiter modifiée avec un pourcentage de patients randomisés mais non analysés, inférieur ou égal à 5% pour Fors 2002.

❖ Concernant le double aveugle :

Il a été considéré comme étant levé si les effets indésirables spécifiques de l'amitriptyline : bouche sèche et ou somnolence (12-13) ont été 2 fois plus fréquents dans le groupe amitriptyline que dans le groupe placebo. Le double aveugle a été levé pour 3 études : Carette 1986, Don Goldenberg 1986 et Ginsberg 1996. Pour ces 3 études, les effets indésirables spécifiques de l'amitriptyline : bouche sèche et somnolence ont été au moins deux fois plus présents dans le groupe amitriptyline que dans les autres groupes de traitement. Pour 6 études, le maintien du double aveugle pose question : elles ne rapportent pas d'information sur les effets

indésirables : Hanonnen 1998 et Fors 2002 (voir ci-dessus), Scudds 1989, Kempenaers 1994, Carette 1995, Braz 2012. Pour 3 études, le double aveugle est maintenu : Carette 1994, Don Goldenberg 1996, Heymann 2001. En effet ces 3 études comparent l'amitriptyline à un placebo et à une molécule active dont les effets indésirables sont proches de ceux de l'amitriptyline, respectivement le cyclobenzaprine, la fluoxétine et la nortriptyline.

❖ Concernant l'analyse de l'étude :

10 études ont une analyse inappropriée. Pour 8 d'entre elles les auteurs ont eu recours à une analyse en intention de traiter modifiée alors que le pourcentage de patients perdus de vue dépasse 5% : Don Goldenberg 1986, Scudds 1989, Carette 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Heymann 2001, Braz 2012. Les 2 autres études sont des analyses per protocole, où des écarts au protocole ont été constatés. De plus, les auteurs n'ont pas contrôlé par une analyse adaptée le respect de l'assignation initiale : Carette 1986 et Kempenaers 1994.

2 études ont une analyse appropriée : Hanonnen 1998 (intention de traiter) et Fors 2002 (intention de traiter modifiée avec 5% de perdus de vue).

Pour Jaeschke 1991, bien que les pourcentages d'effets indésirables n'aient pas été mesurés par les auteurs, la fréquence et la spécificité des effets indésirables de l'amitriptyline, associées à la méthodologie de l'essai : essai ouvert avec amitriptyline suivi par essai cross over en double aveugle, suggèrent fortement que les patients puissent identifier le placebo de la molécule active. De plus l'article rapporte qu'une patiente a réalisé 2 fois l'essai clinique avec des doses différentes d'amitriptyline, ce qui est un écart au protocole de l'étude. Le risque de biais pour ce domaine est donc élevé.

d) Biais résultant des données manquantes

Pour 2 études, le risque de biais lié aux données manquantes est jugé faible. Ainsi, pour Fors 2002 le pourcentage de données manquantes est jugé négligeable (inférieur ou égal à 5%), alors que pour Heymann 2001, les auteurs ont vérifié que les données manquantes (10%) n'avaient pas influencé le résultat final, par un test de Khi 2.

Les 10 autres études (Carette 1986, Don Goldenberg 1986, Scudds 1989, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Braz 2012) sont à risque élevé de biais pour ce domaine, car présentant plus de 5% de données manquantes, jugées probablement en lien avec la vraie valeur du résultat. Parmi ces études, 1 seule (Hanonnen 1998) utilise une méthode pour limiter les données manquantes : le report de la dernière valeur mesurée pour remplacer les mesures ultérieures non réalisées. Cependant cette méthode n'est pas considérée comme valide.

A titre indicatif pour Jaeschke 1991, le pourcentage de données manquantes est nul (pas de perdu de vue), l'étude est donc à faible risque de biais pour ce domaine.

e) Biais résultant de la mesure des critères de jugement

Le risque de biais dans ce domaine est jugé faible pour 3 études : Carette 1994, Don Goldenberg 1996, Heymann 2001, et élevé pour les 9 autres.

En effet la méthode de mesure des critères de jugement pour les 12 études est adaptée. Cependant les critères de jugement de ces 12 études sont subjectifs, et l'évaluateur est toujours le patient avec parfois le médecin examinateur. Or comme nous l'avons vu ci-dessus, le double aveugle est maintenu dans seulement 3 études : Carette 1994, Don Goldenberg 1996 et Heymann 2001. Pour les 9 autres études, le double aveugle a soit été jugé levé (Carette 1986, Don Goldenberg 1986, Ginsberg 1996), soit les éléments disponibles dans les articles ne permettent pas de statuer sur sa levée (Scudds 1989, Kempenaers 1994, Carette 1995, Hanonnen 1998, Fors 2002, Braz 2012), ce qui conduit à un risque de biais jugé élevé.

A titre indicatif, ce domaine de biais est jugé élevé car le double aveugle est considéré comme levé dans l'étude de Jaeschke 1991.

f) Biais résultant de la sélection des résultats rapportés

Aucune des 12 études n'a à priori enregistré son protocole de recherche. Les auteurs n'ont pas donné d'information permettant de savoir si les données ont été analysées conformément à un plan d'analyse prédéfini. 9 études rapportent l'ensemble des résultats des analyses réalisées et ont donc un risque de biais intermédiaire pour ce domaine : Carette 1986, Don Goldenberg

1986, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Heymann 2001, Fors 2002, Braz 2012. 3 études sélectionnent les résultats des analyses réalisées dans leur article : Scudds 1989, Carette 1994, et Kempenaers 1994, et ont donc un risque de biais élevé pour ce domaine.

A titre indicatif, l'étude Jaeschke 1991 n'a pas non plus enregistré de protocole de recherche, et ne rapporte pas les résultats des critères de jugement secondaires pour les périodes de traitement par placebo. Le niveau de risque est donc jugé élevé pour ce domaine.

g) Risque lié à la multiplicité des tests statistiques

Dans 5 études, les auteurs n'ont soit pas défini de critère de jugement principal : Don Goldenberg 1986, Kempenaers 1991, Braz 20012, soit en ont défini plusieurs : Carette 1986 et Fors 2002, ce qui entraîne une inflation du risque alpha.

Dans 7 études, les auteurs ont défini un seul critère de jugement principal : Scudds 1989, Carette 1994, Carette 1995, Ginsberg 1996, Don Goldenberg 1996, Hanonnen 1998, Heymann 2001. Parmi ces 7 études, 4 ont pour critère de jugement principal un score composite ayant nécessité des mesures différentes, ce qui a entraîné une inflation du risque alpha : Scudds 1986, Carette 1994, Carette 1995, Ginsberg 1996. Par ailleurs, considérés indépendamment les uns des autres, les critères de jugement composant un score composite sont considérés comme étant des critères de jugement secondaires, que les auteurs aient réalisé ou non un ajustement du risque alpha. Pour les études Don Goldenberg 1996 et Heymann 2001, les auteurs ne définissent pas explicitement leur critère de jugement principal, mais stipulent avoir utilisé un critère de jugement en particulier pour le calcul de leur effectif. Ces critères de jugement sont donc leur critère de jugement principal. L'étude Hanonnen 1998 définit explicitement un critère de jugement principal unique.

Dans 2 études (Kempenaers 1994 et Ginsberg 1996), les auteurs n'ont pas comparé les résultats finaux des patients traités par amitriptyline et par placebo, mais leurs améliorations respectives. 1 étude n'est pas claire quant à la réalisation d'une comparaison entre les résultats finaux sous traitement par amitriptyline et sous traitement par placebo : Don Goldenberg 1996. Dans le doute, nous avons considéré que les auteurs n'ont pas comparé les résultats finaux (voir chapitre « résumé »). Ainsi, pour ces 3 études, les résultats ne peuvent pas être considérés comme confirmatoires.

Dans 4 études, les auteurs ont utilisé une méthode d'ajustement du risque alpha :

- Don Goldenberg 1986 par la méthode Duncan's
- Scudds 1989 et Carette 1995 par la méthode Tukey's
- Hanonnen 1998 par la méthode de Bonferroni.

Cependant pour Don Goldenberg 1986, les auteurs ne définissent pas de critère de jugement principal, donc malgré la technique d'ajustement utilisée, on considère que les résultats ci-après sont exploratoires : évaluation globale par le patient et par le médecin, douleur, sommeil, fatigue.

Ainsi, si les résultats des études ci-dessous sont significatifs, alors ils sont confirmatoires :

- Scudds 1989 : pour son score composite de la douleur, mais pas pour l'intensité de la douleur
- Carette 1995 : pour son score composite de symptômes, mais pas pour les critères de jugement qui le composent pris indépendamment, c'est-à-dire pas pour la douleur, ni le sommeil, ni la fatigue, ni l'amélioration globale du patient évaluée par ce dernier ou par le médecin
- Hanonnen 1998 : pour son échelle d'évaluation de l'amélioration des patients par le médecin : Clinical Impression of Changment (CIC) (Impression Clinique de Changement)
- Heymann 2001 : pour son critère de jugement principal : FIQ

Les autres résultats sont exploratoires.

A noter que tous les essais cliniques inclus étant à risque de biais global élevé, même s'ils ont des résultats confirmatoires, ces résultats ne peuvent pas être concluants.

A titre indicatif, l'étude Jaeschke 1991 définit deux critères de jugement principaux : un score composite, et le nombre de points sensibles. Elle n'utilise pas de méthode d'ajustement du risque alpha. Ses résultats sont donc exploratoires.

3. Analyse de sensibilité:

Nous avons vu lors de la partie méthodologie que l'outil RoB 2 permet une analyse robuste et reproductible des articles, mais laisse une place d'interprétation personnelle dans certains de ses items. Ceci nous a conduit à statuer sur 3 points avec l'accord de Pr Boussageon :

- **Item 1** : La levée du double aveugle en cas d'absence d'information sur les effets indésirables spécifiques de l'amitriptyline
- **Item 2** : Un seuil de données manquantes et d'acceptabilité de perdus de vue en analyse en intention de traiter modifiée (ITTm) à 5%
- **Item 3** : Une réponse " absence d'information " si les auteurs ne déclarent pas avoir enregistré leur protocole au préalable

Etant donné que ces décisions impactent fortement l'évaluation du risque de biais de chaque catégorie, il a été décidé avec le Pr Boussageon de réaliser une étude de sensibilité méthodologique lorsque cela est pertinent.

Ainsi, pour chaque article scientifique analysé, nous avons regardé si son niveau global de preuve est impacté par l'une de ces trois décisions.

Si c'est le cas, l'article a été réanalysé avec :

- **item 1 réinterprété** : Pour les cas où les effets indésirables ne sont pas mentionnés, une décision d'absence d'impact sur le double aveugle
- **item 2 réinterprété** : Pour les données manquantes et le pourcentage de perdus de vue acceptable lors d'une analyse en intention de traiter modifiée, deux seuils ont été fixés : 10% et 15%.
- **item 3 réinterprété** : En cas d'absence d'information sur l'enregistrement du protocole, l'estimation que cet élément ne constitue pas un biais.

A - S. Carette 1986

L'item 1 réinterprété ne s'applique pas car les auteurs rapportent les effets indésirables, donc le double aveugle reste levé. Par conséquent, le domaine 4 « mesure des résultats » reste à risque de biais élevé.

L'item 2 réinterprété, ne s'applique pas car le pourcentage de perdus de vue est à 15.7%. Donc les domaines 2 « écarts aux interventions », et 3 « données manquantes » restent à risque de biais élevé.

En appliquant l'item 3 réinterprété, le domaine de jugement 5 « sélection des résultats » devient à risque faible.

Impact sur le risque global : Dans tous les cas, le risque global reste élevé car les domaines 2, 3 et 4 restent à risque élevé de biais.

B - Don. Goldenberg 1986

L'item 1 réinterprété ne s'applique pas car les auteurs rapportent les effets indésirables, donc le double aveugle reste levé. Par conséquent, le domaine 4 « mesure des résultats » reste à risque de biais élevé.

En appliquant l'item 2 réinterprété avec un seuil à 10% le domaine 2 passe en faible risque ainsi que le domaine 3.

En appliquant l'item 3 réinterprété, le domaine 5 passe à faible risque.

Impact sur le risque global : dans tous les cas, le risque global reste élevé car le domaine 4 reste à risque élevé de biais.

C - Ra. Scudds 1989

En appliquant l'item 1 réinterprété le domaine 4 passe en risque faible.

En appliquant l'item 2 réinterprété avec un seuil à 10 %, les domaines 2 et 3 passent à risque faible

En appliquant l'item 3 réinterprété, le domaine 5 reste à risque élevé car les auteurs ont sélectionné les résultats rapportés.

Impact sur le risque global : dans tous les cas, le risque global reste élevé car le domaine 5 reste à risque élevé de biais.

D - S. Carette 1994

L'item 1 réinterprété ne s'applique pas car le double aveugle n'était pas levé et que le domaine 4 était déjà à risque faible.

L'item 2 réinterprété, s'applique uniquement au domaine 2 avec le seuil à 15% car le pourcentage de perdus de vue est de 11.5%. La réinterprétation de l'item ne modifie pas l'évaluation du risque du domaine 3 pour lequel le pourcentage de données manquantes de 25% reste élevé.

En appliquant l'item 3 réinterprété, le domaine 5 reste à risque élevé car les auteurs ont sélectionné les résultats rapportés.

Impact sur le risque global : dans tous les cas, le risque global reste élevé car les domaines 3 et 5 restent à risque élevé de biais.

E - C. Kempnaers 1994

En appliquant l'item 1 réinterprété le domaine 4 passe en risque faible de biais.

L'item 2 réinterprété, ne s'applique pas car le pourcentage de perdus de vue est à 36%. Donc les domaines 2 « écarts aux interventions », et 3 « données manquantes » restent à risque de biais élevé.

En appliquant l'item 3 réinterprété, le domaine 5 reste à risque élevé car les auteurs ont sélectionné les résultats rapportés.

Impact sur le risque global : dans tous les cas, le risque global reste élevé (tous les domaines étant en risque élevé hormis le domaine 4 si application de l'item 1)

F - S. Carette 1995

En appliquant l'item 1 réinterprété le domaine 4 passe en risque faible.

En appliquant l'item 2 réinterprété avec un seuil à 10 %, les domaines 2 et 3 passent à risque faible

En appliquant l'item 3 réinterprété, le domaine 5 passe à faible risque également.

Impact sur le risque global : Si application d'un seul item le risque global reste élevé quel que soit l'item. Si application des 3 items alors le risque global devient faible.

G - Don. Goldenberg 1996

L'item 1 réinterprété ne s'applique pas car le double aveugle n'était pas levé et le domaine 4 était déjà à risque faible.

L'item 2 réinterprété, ne s'applique pas car le pourcentage de perdus de vue est à 38.7%. Donc les domaines 2 « écarts aux interventions », et 3 « données manquantes » restent à risque de biais élevé.

En appliquant l'item 3 réinterprété, le domaine 5 passe à faible risque.

Impact sur le risque global : dans tous les cas, le risque global de biais reste élevé car les domaines 2 et 3 sont à risque élevé de biais.

H - F. Ginsberg 1996

L'item 1 réinterprété ne s'applique pas car les auteurs rapportent les effets indésirables, donc le double aveugle reste levé. Par conséquent, le domaine 4 « mesure des résultats » reste à risque de biais élevé.

En appliquant l'item 2 réinterprété avec un seuil à 10 %, les domaines 2 et 3 passent à risque faible

En appliquant l'item 3 réinterprété, le domaine de biais 5 « sélection des résultats » devient à risque faible.

Impact sur le risque global : Dans tous les cas, le risque global reste élevé car le domaine 4 reste à risque élevé

I - P. Hanonnen 1998

En appliquant l'item 1 réinterprété le domaine 4 passe en risque faible.

L'item 2 réinterprété, ne s'applique pas car le domaine 2 utilise une analyse en intention de traité et est donc à risque faible de biais et que le pourcentage de données manquantes est de 29%, donc le domaine 3 « données manquantes » reste à risque de biais élevé.

En appliquant l'item 3 réinterprété, le domaine de jugement 5 « sélection des résultats » devient à risque faible.

Impact sur le risque global : dans tous les cas, le risque global reste élevé car le domaine 3 est à risque élevé

J - R.E Heymann 2001

L'item 1 réinterprété ne s'applique pas car le double aveugle n'était pas levé et le domaine 4 était déjà à risque faible.

En appliquant l'item 2 réinterprété avec un seuil à 10 %, le domaine 2 passe à risque faible. Le domaine 3 est inchangé et reste en risque faible car les auteurs ont vérifié que les perdus de vue n'impactaient pas leur résultat.

En appliquant l'item 3 réinterprété, le domaine de jugement 5 « sélection des résultats » devient à risque faible.

Impact sur le risque global : Si application uniquement de l'item 2 alors le risque global estimé est « quelques inquiétudes » (liées au domaine 5). Si application uniquement de l'item 3 révisé, le risque de biais global reste élevé (biais élevé domaine 2). Si application des items 2 et 3, alors le risque global de biais estimé est faible

K - A Fors 2002

En appliquant l'item 1 réinterprété le domaine 4 passe en risque faible.

L'item 2 réinterprété ne s'applique pas car le pourcentage de perdus de vue est faible : 5% donc les domaines 2 et 3 sont à risques de biais faibles.

En appliquant l'item 3 réinterprété, le domaine de jugement 5 « sélection des résultats » devient à risque faible.

Impact sur le risque global : Si application uniquement item 1, alors le risque global : estimé est « quelques inquiétudes ». Si application uniquement de l’item 3, alors le risque global est « élevé » du fait du domaine 4. Si application des deux items, alors le risque global estimé est faible

M - A. Braz 2012

En appliquant l’item 1 réinterprété le domaine 4 passe en risque faible.

L’item 2 réinterprété, ne s’applique pas car le pourcentage de perdus de vue est à 27%. Donc les domaines 2 « écarts aux interventions », et 3 « données manquantes » restent à risque de biais élevé.

En appliquant l’item 3 réinterprété, le domaine de jugement 5 « sélection des résultats » devient à risque faible.

Impact sur le risque global : dans tous les cas, le risque global reste élevé car les domaines 2 et 3 restent à risque élevé de biais.

A titre indicatif : R. Jaeschke 1991

En appliquant l’item 1 réinterprété le domaine 4 passe en risque faible.

L’item 2 réinterprété ne s’applique pas car il n’y a pas de perdu de vue. Le domaine de biais « données manquantes » était déjà à risque « faible ». Par ailleurs, le domaine de biais 2 reste à risque « élevé » car les auteurs réalisent un écart au protocole susceptible d’avoir affecté le résultat.

En appliquant l’item 3 réinterprété, le domaine 5 reste à risque élevé car les auteurs ont sélectionné les résultats rapportés.

Impact sur le risque global : dans tous les cas, le risque global reste élevé car l’estimation du risque de biais du domaine 1 (processus de randomisation) reste « quelques inquiétudes » et des domaines 2 et 5 restent à risque « élevé ».

En conclusion :

L'analyse de sensibilité méthodologique a été réalisée sur les 12 études incluses dans la méta-analyse ainsi que Jaeschke 1991 (voir tableau 2)

Nom l'auteur principal et année de publication	Biais résultant du processus de randomisation	Biais dû à l'effet de période et à l'effet carry over	Biais résultant du aux écarts aux interventions prévues	Biais résultant des données manquantes	Biais résultant de la mesure du critère de jugement	Biais résultant de la sélection des résultats rapportés	Biais résultant de l'inflation du risque alpha	Evaluation globale Sans prise en compte inflation risque alpha
Carette S. 1986								Risque élevé
Don. Goldenberg 1986								Risque élevé
RA. Scudds 1989								Risque élevé
R. Jaeschke 1991								Risque élevé
S. Carette 1994								Risque élevé
C. Kempenaers 1994								Risque élevé
S. Carette 1995								Risque faible
Don Goldenberg 1996								Risque élevé
F. Ginsberg 1996								Risque élevé
P. Hanonnen 1998								Risque élevé
R.E Heymann 2001								Risque faible
E.A Fors 2002								Risque faible
A. Braz 2012								Risque élevé

Tableau 2 : Représentation du niveau de risque de biais par domaine sur les 12 études incluses et Jaeschke 1991 après nouvelle interprétation des items du RoB 2.

L'étude de sensibilité montre que, même après utilisation de l'outil RoB 2 avec interprétation révisée des 3 items, l'évaluation du risque global de biais des 12 études incluses change peu. Ainsi, en ne modifiant qu'un seul item, par étude, aucune étude ne devient à risque faible de biais. Si on applique au moins deux items révisés, alors 25% des études (Fors 2002, Heymann 2001 et Carette 1995) deviennent à risque de biais global « faible » selon le RoB 2 et 16.67% avec la méthodologie REB (prise en compte du risque alpha : Carette 1995 et Heymann 2001). Les changements d'interprétation impactent principalement les domaines de biais : déviation par rapport aux interventions prévues, mesure du critère de jugement et dans une moindre mesure la sélection dans les résultats rapportés et dans les données manquantes (voir Figure 3).

Pour la suite de notre analyse, nous resterons sur l'évaluation non réinterprétée du risque de biais étant donné qu'il n'était prévu de ne réinterpréter au maximum qu'un seul item par étude.



Figure 3 : Evolution du pourcentage d'études incluses dans la méta-analyse présentant un risque de biais estimé « faible » ou présentant « quelques inquiétudes » ou « élevé », par domaine de biais, selon les 2 interprétations des items du RoB 2 (initiale et révisée)

4. Résultats de chaque étude

Les 12 études incluses dans la méta-analyse et Jaeschke 1991 comparent toujours à minima l'amitriptyline à son placebo. Pour plus de clarté, seuls les résultats concernant l'amitriptyline contre son placebo et les critères de jugement ci-après, seront rapportés dans ce chapitre :

- La douleur
- Les troubles du sommeil
- L'amélioration globale des symptômes évaluée par le patient
- L'amélioration globale des symptômes évaluée par le médecin
- La fatigue
- Les effets indésirables et les perdus de vue

A noter que comme toutes les études sont à haut niveau de biais, aucun résultat n'est concluant selon le REB.

Un résultat a été défini comme étant confirmatoire s'il est la mesure du critère de jugement principal (ou des critères de jugement principaux avec un ajustement du risque alpha) et qu'il est significatif.

Les résultats des 7 études ci-après sont présentés sous forme de graphiques, de pourcentages de patients améliorés ou de fonctions mathématiques : Carette 1986, Don Goldenberg 1986, Scudds 1989, Carette 1994, Kempenaers 1994, Heymann 2001 et Braz 2012. Or seuls les résultats sous forme de moyennes et écarts-types peuvent être intégrés dans l'analyse quantitative qui sera secondairement réalisée. Les auteurs de ces études ont été contactés, mais soit les données des essais cliniques ne sont plus disponibles, soit ils ne nous ont pas répondu. Grâce à l'aide de Mme Veloso Mélanie, biostatisticienne du pôle recherche du Centre Hospitalo-Universitaire de Grenoble, nous avons pu récupérer une partie des données manquantes nécessaires aux analyses qualitatives et quantitatives (estimation des résultats, et des écarts types lorsqu'ils sont représentés, à partir des graphiques).

a) **Carette 1986**

L'étude **Carette 1986** évalue l'efficacité de l'amitriptyline 50 mg par rapport à celle d'un placebo dans la prise en charge de la fibromyalgie sur 70 patients, à l'issue de 9 semaines de traitements. Les critères de jugement principaux sont le pourcentage de patients ayant présenté une amélioration modérée ou marquée de leurs symptômes et / ou du TMS de 8 kg ou plus. Les critères de jugement secondaires sont l'évaluation de la raideur matinale, la douleur, le sommeil et l'évaluation globale des symptômes par le médecin. La douleur est mesurée par EVA (avec une probable erreur de frappe des auteurs qui notent 10 = tolérable, au lieu probablement d'intolérable pour la douleur). Le sommeil est noté par une auto-évaluation avec au choix : plus reposé au matin (1), moins reposé au matin (2), inchangé (3). Le résultat est rapporté sous forme de pourcentage de patients améliorés. La mesure de l'amélioration des symptômes est réalisée sur une échelle discontinue de 1 « pire », 2 « inchangée », 3 « amélioration minime », 4 « amélioration modérée » et 5 « amélioration marquée ». Les effets indésirables ont aussi été mesurés.

Les auteurs n'ont recours à aucune technique d'ajustement du risque alpha. Les résultats de l'étude sont donc exploratoires.

L'étude constate pour l'un des critères de jugement principaux : « pourcentage de patients ayant présenté une amélioration modérée ou marquée de leurs symptômes » une différence significative en faveur de l'amitriptyline pour le résultat intermédiaire à la 5^{ème} semaine et à la fin de l'essai : 9^{ème} semaine :

- 9^{ème} semaine : Amitriptyline 17 patients = 63%, vs placebo 10 patients = 32%, $X^2 = 4.72$, $p = 0.03$.
- 5^{ème} semaine : Amitriptyline 15 patients = 55% vs placebo 7 patients = 22%, $X^2 = 5.73$, $p = 0.002$.

Mais lorsque les auteurs comparent l'amélioration globale des patients (minimale, modérée et marquée), la différence entre le nombre de patients améliorés dans le groupe amitriptyline et le groupe placebo est non significative à la fin du traitement, bien qu'en faveur du traitement actif.

- 5^{ème} semaine : Amitriptyline 21 patients = 77% vs placebo 14 patients = 43%, $X^2 = 7.26$, $p = 0.008$
- 9^{ème} semaine : Amitriptyline 19 patients = 70% vs placebo 16 patients = 50%, $X^2 = 2.55$, $p = 0.12$

Pour les critères de jugement secondaires à la 9^{ème} semaine :

- « Douleur » : les résultats montrent une amélioration non significative dans le groupe amitriptyline : amitriptyline 4.3 ± 3.0 et placebo 5.0 ± 3.0 , p non rapporté
- « Amélioration de la qualité du sommeil » : les résultats montrent un pourcentage significativement plus élevé de patients améliorés dans le groupe traité par amitriptyline par rapport au groupe placebo, respectivement : Amitriptyline 70% vs placebo 40% $X^2 = 5.32$, $p = 0.02$ (avec un résultat déjà significatif à la 5^{ème} semaine de traitement : Amitriptyline 70% vs placebo 34%, $X^2 = 7.26$, $p = 0.008$).
- « Evaluation globale des symptômes par le médecin » : les auteurs n'ont pas rapporté si leurs résultats sont significatifs ou non, ni pour l'amélioration globale (minime, modérée et marquée) ni en ne tenant compte que de l'amélioration modérée et marquée.
- Amélioration globale en faveur de l'amitriptyline à la 5^{ème} et 9^{ème} semaine :
 - 9^{ème}S : amitriptyline 18 patients améliorés sur 27 versus placebo : 14 patients améliorés sur 32
 - 5^{ème}S : amitriptyline 19 patients améliorés sur 27 versus placebo : 12 patients améliorés sur 32

- Amélioration « modérée » ou « marquée » en faveur de l'amitriptyline à la 5^{ème} et la 9^{ème} semaine :
 - 9^{ème}S : amitriptyline 15 patients améliorés sur 27 versus placebo 8 patients améliorés sur 32
 - 5^{ème}S : amitriptyline 11 patients améliorés sur 27 versus placebo 5 patients améliorés sur 32

L'étude rapporte aussi le pourcentage de patients ressentant des effets indésirables : 70% dans le groupe amitriptyline soit 19 patients sur 27 et 12% dans le groupe placebo soit 4 patients sur 32, sans préciser si la différence observée est significative ou non. Les auteurs précisent uniquement la nature des effets indésirables les plus fréquents et non leur répartition (bouche sèche et somnolence).

b) Don Goldenberg 1986

L'étude **Don Goldenberg 1986** évalue l'efficacité de l'amitriptyline 25 mg, du naproxène 500 mg, et de leur association par rapport à celle d'un placebo dans la prise en charge de la fibromyalgie sur 62 patients, à l'issue de 6 semaines de traitement. Les auteurs n'ont pas défini de critères de jugement principaux ni secondaires. Les critères de jugement sont l'évaluation globale des symptômes par le patient et le médecin, la raideur matinale, la douleur, la fatigue, le sommeil et le TMS. Hormis le TMS, ils sont mesurés par l'EVA (0 = asymptomatique, 10 = très symptomatique)

Les auteurs ne définissent pas de critère de jugement principal, donc bien qu'ils utilisent la méthode de Duncan pour réaliser un ajustement du risque alpha, les résultats restent exploratoires.

Les résultats sont présentés sous forme de graphiques. Ils montrent une différence significative en faveur du groupe traité par amitriptyline pour les critères de jugement (estimation des moyennes rapportées ci-dessous, calcul ou estimation des écarts types impossible) :

- Fatigue : Amitriptyline 4.3 placebo 7.5 ($p < 0.05$)
- Sommeil : Amitriptyline 3.0 placebo 6.7 ($p < 0.05$)
- Evaluation globale par le patient Amitriptyline 4.5 placebo 7.9 ($p < 0.05$)

- Evaluation globale par le médecin Amitriptyline 4.1 placebo 8.8 ($p < 0.01$)
- Douleur : Amitriptyline 5.4 placebo 7.3 ($p < 0.01$).

Les auteurs rapportent aussi dans leur article le constat d'une amélioration significative intragroupe pour l'amitriptyline pour les critères de jugement : douleur, sommeil, fatigue, évaluation globale par le médecin et le patient ($p < 0.001$).

Concernant les effets indésirables, les auteurs rapportent avoir eu 8 patients atteints d'effets indésirables. 4 sont spécifiques de l'amitriptyline : bouche sèche, et sont présents chez des patients ayant reçu de l'amitriptyline (amitriptyline + placebo ou de l'amitriptyline + naproxène). 4 (2 diarrhées et 2 dyspepsies) sont non spécifiques de l'amitriptyline et ont été rapportés par des patients non traités par amitriptyline (naproxène + placebo ou placebo seul).

c) Scudds 1989

L'étude **Scudds 1989** évalue l'efficacité de l'amitriptyline 50 mg par rapport à celle d'un placebo dans la prise en charge de la fibromyalgie sur 39 patients, à l'issue d'une période de 4 semaines de traitement (essai en cross-over). Les auteurs ont défini comme critères de jugement principal un score composite évaluant la douleur par l'évaluation du TMS, le seuil de douleur, la tolérance de la douleur, et l'intensité de la douleur. L'intensité de la douleur est mesurée par le questionnaire de Mc Gill. Un critère de jugement secondaire est aussi exploré : le nombre de patients présentant une amélioration globale de leurs symptômes. L'amélioration des symptômes est mesurée sur une échelle discontinue de 1 à 5 précédemment décrite dans l'étude de Carette 1986. L'étude rapporte aussi une analyse du sous-groupe composé uniquement des patients ayant eu une évolution globale des symptômes positive « modérée » ou « marquée » du critère de jugement principal (score composite évaluant la douleur).

Le critère de jugement principal de l'étude est un score composite mesurant différents aspects de la douleur, mais seule l'intensité de la douleur est l'un de nos critères de jugement. Mesurée indépendamment du score composite, elle est un critère de jugement secondaire et est donc exploratoire. Malgré le fait que les auteurs utilisent une méthode d'ajustement du risque alpha (méthode de Tukey), les résultats de nos critères de jugement sont exploratoires.

Les résultats de l'étude montrent pour le critère de jugement principal : score composite évaluant la douleur, une différence significative en faveur de l'amitriptyline uniquement pour

deux des quatre critères de jugement : score total de douleur musculaire (TMS) (HSD = 3.74, $p < 0.05$) et intensité de la douleur (HSD = 3.54, $p < 0.05$). Les auteurs n'ont pas réalisé de comparaison entre les périodes de traitement par placebo et celles par amitriptyline pour les deux autres critères de jugement du score composite : seuil et tolérance de la douleur. En effet, leurs résultats intragroupes ne mettent pas en évidence de différence significative au cours du temps.

Les résultats de l'étude montrent pour le critère de jugement secondaire : nombre de patients présentant une amélioration globale de ces symptômes jugée « minime » ou « modérée » ou « marquée », une différence significative en faveur de l'amitriptyline (Amitriptyline 27 patients sur 36 (75%) et placebo 8 patients sur 37 (22%) $X^2 = 21.6$, $p < 0.001$).

Les résultats de l'analyse du sous-groupe (patients ayant eu une amélioration « modérée » ou « marquée » de leurs symptômes) trouvent une différence significative en faveur de l'amitriptyline pour le score composite « douleur » ($t = 2.01$, $p < 0.05$) avec une différence significative observée par les critères de jugement « Score total de douleur musculaire : TMS » (HSD = 4.97, $p < 0.01$), intensité de la douleur (HSD = 4.36, $p < 0.01$), et seuil de douleur (HSD = 4.34, $p < 0.05$). Les auteurs ne rapportent pas dans l'article la comparaison pour la tolérance de la douleur, car leur analyse du sous-groupe ne met pas en évidence de différence significative pour ce critère de jugement au cours du temps.

d) Carette 1994

L'étude **Carette 1994** évalue l'efficacité de l'amitriptyline 50 mg et de la cyclobenzaprine par rapport à celle d'un placebo dans la prise en charge de la fibromyalgie sur 208 patients, à l'issue de 6 mois de traitement. Le critère de jugement principal est un score composite évaluant l'amélioration de divers symptômes : au moins 4 sur 6 avec 50% d'amélioration pour la douleur, le sommeil, la fatigue, l'évaluation globale par le patient et / ou le médecin, et ou la majoration d'1 kg au moins du total myalgique score. Hormis le TMS, les autres critères de jugement du score composite sont mesurés par l'échelle EVA (avec 0 = symptomatique, et 10 : très symptomatique). Les critères de jugement secondaires sont : la douleur, la fatigue, le sommeil, la sensation de repos au réveil, la raideur matinale, et l'évaluation globale par le patient et le médecin, le score total de douleur, les limitations d'activités dues à la pathologie et l'état émotionnel (anxiété / dépression).

Les auteurs ont un critère de jugement unique, mais c'est un score composite ayant nécessité plusieurs mesures. Ils n'ont pas utilisé de méthode pour limiter l'inflation du risque alpha. Les résultats de l'étude sont donc exploratoires.

Les critères de jugement secondaires sont rapportés sous forme de graphique. Ils mettent en évidence une différence non significative en faveur de l'amitriptyline à l'issue des 6 mois de traitement (estimation des moyennes rapportées ci-dessous, calcul ou estimation des écarts types impossible) :

- Douleur : Amitriptyline 47 placebo 52 (p non significatif)
- Sommeil : Amitriptyline 38.5 placebo 50 (p non significatif)
- Fatigue : Amitriptyline 50 placebo 59.5 (p non significatif)
- Evaluation globale des symptômes par le patient : Amitriptyline 48 placebo 52.5 (p non significatif)
- Evaluation globale des symptômes par le médecin : Amitriptyline 40 placebo 48 (p non significatif)

A noter que les résultats intermédiaires comparant l'amitriptyline au placebo mettent en évidence une différence significative en faveur du traitement actif, jusqu'au 1^{er} mois pour la douleur, 4^{ème} mois pour la fatigue et l'évaluation globale des symptômes par le patient et 5^{ème} mois pour le sommeil et l'évaluation globale par le médecin.

Les résultats des autres critères de jugement secondaires comme l'évaluation de la douleur par le score de Mc Gill ne montrent pas de différence significative entre le groupe traité par amitriptyline et celui traité par placebo. Par contre les résultats montrent une amélioration significative intragroupe pour les 3 points du score de Mc Gill pour l'amitriptyline ($p < 0.001$), et pour le placebo ($p < 0.05$) :

- Intensité ($A = 2.17 \pm 1.02$ et $P = 2.47 \pm 0.97$, $p > 0.05$)
- Rating index ($A = 19.5 \pm 13.5$, $P = 21. \pm 14.4$, $p > 0.05$)
- Nombre de mots choisis $A = 8.87 \pm 5.79$, $P = 10.2 \pm 5.62$, $p > 0.05$)

Les auteurs rapportent que 95% de leurs patients traités par amitriptyline et 62% de ceux traités par placebo ont ressenti des effets indésirables. L'article précise qu'ils ne diffèrent pas en termes de nature et de fréquence entre les 2 traitements actifs (cyclobenzaprine et amitriptyline).

e) **Kempenaers 1994**

L'étude **Kempenaers 1994** évalue l'efficacité du SER282 et de l'amitriptyline 50 mg par rapport à celle d'un placebo dans la prise en charge de la fibromyalgie sur 36 patients, à l'issue de 8 semaines de traitement. Les critères de jugement sont le nombre de points sensibles, un score de douleur en des points définis (score de 0 à 18), les douleurs diffuses (score de 0 à 3), l'intensité de la douleur globale par VAS (0 = asymptomatique, 10 = la pire douleur), les troubles du sommeil mesurés par une échelle de 0 à 3, les symptômes associés (score composite regroupant les troubles du sommeil, la fatigue matinale, la fatigue en générale, les crampes, les dysménorrhées, la constipation, les céphalées, les gonflements des doigts, les vertiges). Les auteurs ont aussi mesuré le jour des EEG (ElectroEncéphaloGramme) : la dépression, l'anxiété, la qualité globale du sommeil par VAS (0 = très symptomatique, 100 = asymptomatique), la fatigue diurne mesurée par une échelle de 0 à 3, la sensation de repos au matin. Ils ont observé la physiologie du sommeil par polysomnographie, ainsi que la tolérance et l'efficacité des traitements.

Les auteurs ne définissent pas de critère de jugement principal et ne rapportent pas l'utilisation d'une méthode d'ajustement du risque alpha. Les résultats ci-dessous sont donc exploratoires. Les auteurs ont comparé les améliorations mesurées en intragroupe pour les groupes de traitements : amitriptyline et placebo par une analyse de variance à mesures répétées, en fonction du temps et de l'interaction groupe et temps.

A l'issue de l'étude, ils retrouvent donc une différence significative entre les groupes au cours du temps pour les critères de jugement :

- Sévérité de la douleur localisée : Amitriptyline 9 ± 4 , placebo 6 ± 3 , p temps x groupe = 0.035
- Sévérité de la polyalgie : Amitriptyline 1.7 ± 1.0 , Placebo 1.1 ± 1.0 , p temps x groupe = 0.045
- Fatigue diurne : Amitriptyline 1.2 ± 0.5 , Placebo 2.7 ± 1.5 , p temps x groupe = 0.05

Ils ne retrouvent pas de différence significative pour les autres critères de jugement :

- Intensité de la douleur globale : Amitriptyline 32 ± 31 , placebo 37 ± 28 , p non significatif
- Qualité globale du sommeil : Amitriptyline 47 ± 30 et placebo 60 ± 22 , p non significatif

Concernant le critère de jugement « amélioration globale des symptômes évaluée par le patient et ou le clinicien » les auteurs rapportent dans le texte de l'article mais sans donner les résultats ni la méthode de mesure, une amélioration intragroupe significative à partir de la 4^{ème} semaine pour l'amitriptyline et de la 8^{ème} semaine de traitement pour le placebo, sans différence significative entre les 2 groupes à la fin de l'essai.

f) **Carette 1995**

L'étude **Carette 1995** évalue l'efficacité de l'amitriptyline 25 mg par rapport à celle d'un placebo dans la prise en charge de la fibromyalgie sur 22 patients, à l'issue de 8 semaines de traitement (cross-over). Le critère de jugement principal est une étude de l'impact d'un type d'anomalie du sommeil (alpha Non Rapide Eyes Movement) sur les symptômes de la fibromyalgie (testé par un score composite : amélioration d'au moins 4 sur 6 symptômes avec 50% d'amélioration pour la douleur, le sommeil, la fatigue, l'évaluation globale par le patient et / ou le médecin, et / ou la majoration d'1 kg au moins du score total de douleur musculaire) et la prédiction de réponse au traitement par amitriptyline. Les critères de jugement secondaires sont l'évaluation indépendante de chacun des critères du score composite. Chacun à l'exception du TMS, est mesuré par EVA (0 = asymptomatique, 10 = très symptomatique)

Le critère de jugement principal est un score composite et une partie des scores qui le composent sont nos critères de jugement. Etant donné que les résultats de ces scores sont présentés de façon indépendante du score composite, ils doivent être considérés comme étant des critères de jugement secondaires. Donc malgré le fait que les auteurs utilisent une méthode d'ajustement du risque alpha (méthode de Tukey), les résultats inclus sont exploratoires.

Les résultats des critères de jugement secondaires ci-après mettent en évidence une différence significative en faveur de l'amitriptyline :

- Douleur : Amitriptyline 5.07 (3.22), placebo 7.13 (2.41), $p < 0.05$
- Sommeil : Amitriptyline 3.93 (3.14) Placebo 6.51 (2.69) $p < 0.05$
- Fatigue : Amitriptyline 5.62 (3.07) Placebo 7.64 (1.80) $p < 0.05$
- Evaluation globale des symptômes par le patient : Amitriptyline 5.47 (3.03) Placebo 7.11 (2.14) $p < 0.05$

- Evaluation globale des symptômes par le médecin : Amitriptyline 4.81 (2.81) Placebo 6.36 (1.59) $p < 0.05$

g) **Don Goldenberg 1996**

L'étude **Don Goldenberg 1996** évalue l'efficacité de l'amitriptyline 25 mg, de la fluoxétine 20mg et de leur interaction, par rapport à un placebo dans un essai clinique en cross-over. 31 patients ont participé à l'essai clinique. Les patients ont été suivis au cours de 4 périodes de traitement de 6 semaines chacune (de façon aléatoire : amitriptyline + placebo, fluoxétine + placebo, placebo seul et amitriptyline + fluoxétine). Le critère de jugement principal n'est pas clairement défini dans l'article, mais les auteurs utilisent « l'évaluation globale de la pathologie par le patient » mesuré par EVA pour calculer l'effectif de patients requis. C'est donc le critère de jugement principal. Les critères de jugement secondaires sont : le TMS, l'évaluation des symptômes par le médecin, la douleur, les troubles du sommeil, la fatigue, la sensation de repos au réveil, l'incapacité due à la pathologie et la dépression.

Les résultats des critères de jugement principaux / secondaires ci-après mettent en évidence une différence significative en faveur de l'amitriptyline :

- Douleur : amitriptyline 64.4 (28.3) et placebo 81.5 (16.5), $p^* = 0.02$,
- Sommeil : amitriptyline 57.0 (34.8) et placebo 74.6 (23.9), $p^* < 0.001$,
- Évaluation globale par le patient : amitriptyline 61.6 (29.5) et placebo 76.8 (24.8), $p^* = 0.02$
- Évaluation globale par le médecin examinateur : amitriptyline 64.2 (25.2) et placebo 74.7 (19.9), $p^* = 0.04$
- Incapacité mesurée par le Fibromyalgia Impact Questionnaire (FIQ) : amitriptyline 52.3 (22.9), placebo 58.5 (17.1) $p^* = 0.03$

Le critère de jugement « fatigue » montre une différence non significative en faveur de l'amitriptyline (amitriptyline 67.7 (39.9), placebo 73.7 (25.1) et $p^* = 0.26$)

Selon la légende du tableau 2 de l'article, p^* désigne la significativité après ajustement à la période et à l'effet principal sur le patient. On constate ainsi que pour tous les critères de jugement présentés, le p^* est significatif (à l'exception de la « fatigue »). Les résultats des comparaisons entre les améliorations dans le groupe amitriptyline et dans le groupe placebo ne

sont pas rapportés dans les tableaux de résultats, mais précisés dans le texte de l'article. Ainsi, les auteurs notent « There was statistically significant improvement, compared with the placebo-treated group, for both the AM and FL treated groups, as measured by FIQ and by the VAS for pain, global well being, and sleep[...] Similar, but not statistically significant, trends for both single drug and combination-drug therapy were noted using the BDI scale, physician global VAS, VAS fatigue ». Donc bien que l'amélioration pour l'évaluation par le médecin soit statistiquement significative dans le tableau de résultats ($p = 0.04$), sa comparaison à l'amélioration sous placebo est non significative. Les résultats présentés dans le tableau 2 de l'article et rapportés ci-dessus montrent donc l'amélioration intragroupe de l'amitriptyline pour les critères de jugement douleur, sommeil, Fibromyalgie Impact Questionnaire et évaluation globale par le médecin examinateur et par le patient. De plus nous avons constaté que, pour chacun de ces critères de jugement, une fois les valeurs incluses dans les analyses quantitatives (voir ci-dessous), l'intervalle de confiance contient systématiquement la valeur 0. Ceci suggère que les résultats finaux de l'amitriptyline sont non significatifs par rapport aux résultats finaux obtenus avec le placebo. Compte tenu du manque de clarté sur la comparaison des résultats finaux du groupe amitriptyline et du groupe placebo, nous avons statué, en accord avec le Dr Yves Baro, que les auteurs de l'article n'ont probablement pas comparé les résultats finaux obtenus dans le groupe amitriptyline avec ceux du groupe placebo. Ainsi, bien que les auteurs aient un critère de jugement principal unique, le résultat de ce critère de jugement sera considéré comme exploratoire, comme les autres résultats de l'étude. A noter que les auteurs ne rapportent pas l'usage d'une méthode pour limiter l'inflation du risque alpha.

Concernant les effets indésirables, les auteurs rapportent uniquement ne pas avoir constaté de différence en termes d'effets indésirables entre les 4 périodes de traitement.

h) Ginsberg 1996

L'étude **Ginsberg 1996** évalue l'efficacité de l'amitriptyline à libération prolongée 25 mg à un placebo dans la prise en charge de la fibromyalgie sur 51 patients, à l'issue de 8 semaines de traitement. Le critère de jugement principal est le pourcentage de patients considérés comme répondeurs au traitement. Il est défini par un score composite : amélioration d'au moins 3 critères sur 4 : 50% ou plus sur l'évaluation globale des symptômes par le patient ou le médecin examinateur, 50% sur la douleur, 25% sur le TMS). Les critères de jugement secondaires sont

: l'évaluation globale des symptômes par le patient ou le médecin évaluateur, la douleur, les perturbations du sommeil, la sensation de repos au réveil et la fatigue qui sont mesurés par EVA ; la raideur matinale mesurée en minutes, le nombre de points sensibles et le TMS. Les auteurs ont aussi mesuré la tolérance clinique du traitement.

Les auteurs ont un critère de jugement unique, mais c'est un score composite ayant nécessité plusieurs mesures. Ils n'ont pas utilisé de méthode pour limiter l'inflation du risque alpha. Les résultats de l'étude sont donc exploratoires.

Les auteurs ne rapportent une comparaison du placebo à l'amitriptyline que pour 3 critères : le critère de jugement principal, la sensation de fatigue au réveil et la fatigue (critères de jugement secondaires). Pour ce dernier critère, les résultats montrent une différence significative en faveur du groupe traité par le principe actif uniquement à la 8^{ème} semaine de traitement :

- 8^{ème} semaine : Amitriptyline 58% patients améliorés versus placebo 18%, $p = 0.014$, chi-squared tests.
- 4^{ème} semaine : Amitriptyline 46% de patients améliorés versus placebo 18% $p = 0.094$, chi-squared tests

De plus, un nombre significativement plus important de patients traités par placebo ressentent une fatigue jugée « fréquente ou permanente » par rapport au nombre de patients traités par amitriptyline (amitriptyline 13, placebo 19, $p = 0.04$ chi-squared test) à la 8^{ème} semaine de traitement.

Les auteurs rapportent une amélioration significative de tous les critères de jugement secondaires dans le groupe amitriptyline au cours du temps (intragroupe) alors qu'aucune amélioration significative n'est observée dans le groupe placebo au cours du temps ce qui les conduit à conclure à une amélioration significative du groupe traité par amitriptyline par rapport au groupe placebo (résultats à la 8^{ème} semaine de traitement) :

- Evaluation globale par le patient Amitriptyline 3.9 ± 2.3 ($p < 0.001$), placebo 6.8 ± 1.8 p non significatif
- Evaluation globale ou par le médecin Amitriptyline 3.5 ± 2.1 ($p < 0.001$), placebo 5.9 ± 1.7 p non significatif
- Evaluation de la douleur Amitriptyline 3.8 ± 2.4 ($p < 0.001$), placebo 7.0 ± 1.3 p non significatif

- Evaluation de la fatigue Amitriptyline 3.8 ± 2.5 ($p = 0.001$), placebo 5.9 ± 2.2 p non significatif
- Evaluation des troubles du sommeil Amitriptyline 2.6 ± 3.1 ($p = 0.003$), placebo 5.1 ± 3.0 p non significatif

Le pourcentage de patients ayant ressenti des effets indésirables est significativement ($p = 0.01$) plus important dans le groupe traité par amitriptyline 29% (soit 7 sur 24 patients) que dans le groupe placebo (aucun patient). De plus, 3 patients ont ressenti des effets indésirables typiques de l'amitriptyline : sécheresse buccale. Les autres effets ressentis sont non spécifiques de la molécule.

i) Hannonen 1998

L'étude **Hannonen 1998** évalue l'efficacité du moclobemide (dose entre 450 mg et 600 mg) par rapport au placebo et à l'amitriptyline (dose entre 25 et 37.5 mg selon la tolérance et l'efficacité ressenties par les patients), dans la prise en charge de la fibromyalgie sur 130 patients, à l'issue de 12 semaines de traitement. Le critère de jugement principal est le pourcentage de patients considérés comme répondeurs au traitement (défini par un score de 3, 2 ou 1 sur l'échelle CIC : Impression Clinique de Changement). Les critères de jugement secondaires sont : l'évaluation globale des symptômes par le patient (EVA), l'évaluation de la douleur (EVA), la quantité et la qualité du sommeil (EVA), la fatigue (EVA), l'échelle fonctionnelle de Sheehan, l'échelle de qualité de vie Nottingham Health Profile, l'échelle Impression de Sévérité Clinique par le médecin examinateur CIS, et le nombre de points sensibles.

Les auteurs ont un critère de jugement unique qui est donc confirmatoire. De plus les auteurs ont utilisé une méthode d'ajustement du risque alpha : méthode de Bonferroni. Les autres critères de jugement de l'étude sont des critères de jugement secondaires et donc exploratoires.

Les résultats montrent une amélioration significative du groupe traité par amitriptyline par rapport au placebo pour :

- Le critère de jugement primaire :
 - o Evaluation de l'amélioration du patient par l'échelle CIC : Amitriptyline 1.14 [1.05] placebo 0.38 [1.42] $p = 0.003$,

- Nombre de patients répondeurs selon l'échelle CIC : 74 % groupe amitriptyline et 49% groupe placebo $p = 0,017$
- Le critère de jugement secondaire :
 - « Sommeil » Amitriptyline 3.6 [2.8] placebo 4.8 [2.9] $p < 0.01$

Pour les autres critères de jugement, l'amélioration intragroupe des patients traités par amitriptyline est significative. Cependant, lors de la comparaison du résultat avec l'amélioration mesurée dans le groupe placebo, la différence ne reste pas significative.

- « Évaluation globale par le patient » amitriptyline 4.5 [2.6] placebo 5.3 [2.5]
- « Douleur » amitriptyline 4.5 [2.8] placebo 5.2 [2.7]
- « Fatigue » amitriptyline 4.7 [2.8] placebo 4.6 [2.6]
- « Impression de sévérité globale CIS par le médecin » amitriptyline 3.4 [0.94], placebo 3.78 [1.15]

Les auteurs ont aussi mesuré les effets indésirables et rapportent des pourcentages proches en termes de nombre de patients pour chaque groupe de traitement : amitriptyline 74%, placebo 80% (pas de différence significative). Cependant ils diffèrent en nature : fatigue et bouche sèche pour l'amitriptyline et fatigue et céphalées pour le placebo, sans précision sur la significativité.

j) Heymann 2001

L'étude **Heymann 2001** évalue l'efficacité de l'amitriptyline 25 mg et de la nortriptyline 25 mg par rapport au placebo dans la prise en charge de la fibromyalgie sur 118 patients, à l'issue de 8 semaines de traitement. Les auteurs ne définissent ni critère de jugement principal ni critère de jugement secondaire ; mais ils ont estimé leur effectif de patients à inclure en fonction du critère de jugement « amélioration de la qualité de vie, évaluée par le questionnaire FIQ. ». C'est donc leur critère de jugement principal. Les critères de jugement secondaires sont : le nombre de points sensibles et l'évaluation par le patient de l'amélioration globale de son état de santé (par la Verbal Scale Global Evaluation : VSGI). Les auteurs mesurent aussi les effets indésirables.

Les auteurs ont un critère de jugement unique, mais non statistiquement significatif, le résultat est donc exploratoire. De plus, ils ne rapportent pas l'usage d'une méthode pour limiter l'inflation du risque alpha. Tous les résultats de l'étude sont donc exploratoires.

Pour le critère de jugement « évaluation par le patient de l'amélioration globale de son état de santé », l'article montre une différence significative en faveur du traitement par amitriptyline comparé au traitement par placebo :

- Amitriptyline 31 patients améliorés soit 86.5%, placebo 18 patients améliorés soit 54.5%, $\chi^2 = 6.6$,

$p < 0.00981$

Pour le critère de jugement : « qualité de vie » évaluée par la FIQ, les résultats montrent une amélioration non significative en faveur de l'amitriptyline par rapport au placebo :

- Score FIQ post traitement : amitriptyline 39.97 ± 6.54 , placebo 51.68 ± 7.98 , p non significatif
- Amélioration mesurée dans le groupe amitriptyline : 36.5%, placebo 24%, p non significatif

Les auteurs ont aussi mesuré les effets indésirables et ne retrouvent pas de différence significative entre les trois groupes de traitements ($p = 0.092$) en termes de fréquence et de nature.

L'étude **Fors 2002** évalue l'efficacité de l'amitriptyline 50 mg à un placebo d'une part et à deux différents types d'imagerie guidée versus un groupe contrôle d'autre part, dans la prise en charge de la fibromyalgie sur 63 patients, à l'issue de 4 semaines de traitement. Les critères de jugement principaux sont : en premier la mesure de la douleur pour les thérapeutiques « imagerie », en deuxième la mesure de la douleur pour le traitement pharmacologique, et en troisième la recherche d'une interaction entre les 2 types de traitements précédents. La mesure de la douleur est basée sur un auto-recueil quotidien (journal de bord tenu quotidiennement par le patient). Les critères de jugement secondaires sont : la différence entre la douleur recueillie par le journal et celle enregistrée dans le laboratoire de recherche, la mesure de l'anxiété, de la dépression et des pensées automatiques négatives. Les auteurs recueillent aussi les effets indésirables.

Les auteurs ont trois critères de jugement principaux. Ils n'ont pas utilisé de méthode pour limiter l'inflation du risque alpha. Les résultats de l'étude sont donc exploratoires.

Les résultats du critère de jugement principal « évaluation de la douleur quotidienne par rapport au traitement médicamenteux reçu » ne retrouve pas de différence entre l'amitriptyline et le placebo, que ce soit en intention de traiter ($F(1.1452) = 0.00$ avec $p = 0.98$) ou en per protocole (amitriptyline 22 patients => EVA : 39.2 (29.1), placebo : 27 patients => EVA : 50.7 (24.1), $F(2.1453) = 0.16$ avec $p = 0.85$). Les auteurs ne retrouvent pas non plus de lien entre l'augmentation de la concentration d'amitriptyline dans le sang des patients et l'effet antalgique ressenti ($p = 0.39$). L'étude post hoc réalisée par les auteurs constate qu'ils avaient une puissance de 97% permettant de montrer une différence entre l'amitriptyline et le placebo s'il y en avait une.

Les effets indésirables n'ont pas été rapportés dans l'article.

k) **Braz 2012**

L'étude **Braz 2012** évalue l'efficacité de p. Ginseng (100 mg) par rapport à l'amitriptyline 25 mg et à un placebo dans la prise en charge de la fibromyalgie sur 52 patients, à l'issue de 12 semaines de traitement. Les auteurs n'ont pas défini de critères de jugement principaux ou secondaires. Ils ont mesuré : le nombre de points sensibles, la douleur, la fatigue, la qualité du sommeil, l'anxiété, et la qualité de vie (FIQ).

Les auteurs ne définissent pas de critère de jugement principal et n'utilisent pas de méthode pour limiter l'inflation du risque alpha. Les résultats de l'étude sont donc exploratoires.

L'estimation approximative des résultats à partir du graphique montre une différence non significative en faveur du placebo pour les critères de jugement :

- Douleur : amitriptyline 4.8 [0.8], placebo : 3.9 [0.8] avec ($p > 0.05$)
- Fatigue : amitriptyline : 4.1 [0.9], placebo : 3.7 [0.4] avec ($p > 0.05$)
- Qualité du sommeil : amitriptyline : 2.8 [0.9], placebo : 2.6 [0.8] avec ($p > 0.05$)
- Qualité de vie (FIQ) : amitriptyline : 32 [6], placebo : 32 [4] avec ($p > 0.05$).

A titre indicatif, l'étude *Jaeschke 1991* évalue l'efficacité de l'amitriptyline à une dose définie avec chaque patient, variant de 5 à 50 mg à un placebo, dans la prise en charge de la fibromyalgie sur 23 essais effectués sur 22 patients, à l'issue d'une succession variable de périodes de traitement (essai en N-of-1 RCT). Les auteurs ont défini deux critères de jugement principaux : le nombre de points sensibles et un score composite évaluant les « symptômes standards ». Ce dernier est composé de l'évaluation du manque d'énergie, de la fatigue, des douleurs, de la raideur matinale, des difficultés de sommeil, des céphalées, et des problèmes intestinaux. 3 critères de jugement secondaires ont aussi été explorés : l'impact de l'essai N-of-1 RCT sur la conduite à tenir du médecin, la confiance du médecin dans la conduite à tenir décidée et les résultats de chaque symptôme composant le score composite. Chacun de ces critères de jugement est évalué par une échelle discontinue de 1 = très symptomatique à 7 = asymptomatique.

L'article rapporte pour les critères de jugement « douleur » et « sommeil » un bénéfice significatif de l'amitriptyline chez 6 / 23 patients, et 3 patients / 23 pour la fatigue ($p < 0.05$). Le nombre de patients ayant présenté par critère de jugement un bénéfice au placebo, n'est pas rapporté. Ces données manquantes empêchent l'inclusion de l'étude dans l'analyse qualitative et dans l'analyse quantitative.

5. Synthèse des résultats :

La synthèse des résultats vise à vérifier par analyse qualitative et quantitative des 12 études incluses, l'efficacité de l'amitriptyline, à travers 6 critères de jugement que nous jugeons pertinents pour les patients :

L'amélioration globale des symptômes évaluée par le patient

- L'amélioration globale des symptômes évaluée par le médecin
- La douleur
- Les troubles du sommeil
- La fatigue
- Les effets indésirables et les perdus de vue

La taille d'effet (Standardized Mean Deviation (SMD) (différence de moyennes standardisées) ou Risk Ratio) est mesurée avec l'outil RevMan et interprétée par le I^2 de Hedges (ajusté) avec I^2 entre 0,2–0,5 = taille d'effet faible, I^2 entre 0,5–0,8 = taille d'effet modéré et $I^2 > 0,8$ = grande taille d'effet.

Les 12 études étant à niveau de risque élevé selon l'outil RoB 2, une méta-analyse ne sélectionnant que les résultats concluants, c'est-à-dire les résultats confirmatoires issus d'essais cliniques de bonne qualité méthodologique, n'a pas pu être réalisée selon la méthodologie REB (14). L'analyse quantitative de sensibilité des résultats a été réalisée sur 11 études parmi les 12 incluses, selon la méthode GRADE (15).

En effet, comme nous l'avons vu dans le chapitre précédent, les résultats des articles ci-après ne sont pas présentés sous forme de moyennes avec écarts types mais sous forme de graphiques, de pourcentages de patients améliorés ou de fonctions mathématiques. De ce fait, les résultats des essais cliniques ci-dessous n'ont pas pu ou n'ont pu être que partiellement inclus dans les analyses qualitatives et quantitatives :

Aucun résultat de l'étude ci-dessous n'a pu être inclus dans l'analyse quantitative :

- Don Goldenberg 1986 :

Les résultats des différents critères de jugement sont uniquement rapportés sous forme de graphique pour le groupe placebo. Une approximation des résultats a pu être réalisée en analysant le graphique. Cependant, les écarts types n'ont pu être ni estimés ni calculés. De plus l'article ne précise pas l'effectif de ses 4 groupes de traitements : amitriptyline + placebo,

amitriptyline + naproxène, naproxène + placebo et placebo. Pour ces différentes raisons, l'article n'a pas été inclus dans l'analyse quantitative. Cependant sachant qu'il y a 62 patients randomisés, avec 1 perdu de vue par groupe et que la répartition est décrite comme réalisée par « une méthode qui assure un assignement équilibré », nous pouvons estimer qu'il doit y avoir environ 15 patients dans le groupe amitriptyline + placebo et 15 patients dans le groupe placebo. Nous utiliserons ces valeurs approximatives lors des calculs d'effectifs.

Une partie des résultats des 4 études ci-dessous n'ont pas pu être inclus dans l'analyse quantitative :

- Carette 1986 :

Seuls les résultats du critère de jugement « intensité de la douleur » sont présentés sous forme de moyennes avec écart type. Pour les critères de jugement « amélioration globale évaluée par le patient » « qualité du sommeil », « évaluation globale des symptômes par le médecin » les auteurs présentent leurs résultats sous forme de nombre de patients améliorés et pourcentage de patients améliorés et ne peuvent donc pas être inclus dans l'analyse quantitative.

- Scudds 1989 :

Les résultats de « l'intensité de la douleur » et de « l'amélioration des symptômes » sont présentés sous forme de fonction numérique ou de pourcentage de patients améliorés. Il n'a pas été possible de calculer à partir de ces données les moyennes et écarts types de nos critères de jugement. Ces résultats n'ont donc pas été inclus dans la méta-analyse. L'essai clinique a pu être inclus pour l'analyse quantitative du nombre de perdus de vue.

- Carette 1994 :

Seuls les résultats du critère de jugement « intensité de la douleur » mesurés par le score de McGill ont été inclus dans l'analyse quantitative. En effet, les résultats des critères de jugement secondaires mesurés par EVA « douleur », « fatigue », « amélioration globale par le patient », « amélioration globale par le médecin » et « troubles du sommeil » ont été rapportés dans l'article sous forme de graphiques. Une approximation des résultats a été estimée en analysant le graphique. Cependant, les écarts types n'ont pu être ni estimés ni calculés. Ces données n'ont donc pas été incluses dans la méta-analyse de sensibilité.

- Kempenaers 1994 :

Les résultats des critères de jugement « amélioration globale des symptômes évaluée par le patient et / ou le clinicien » ne sont pas rapportés dans l'article. Seule l'absence de différence significative entre les 2 groupes est mentionnée. Les résultats pour l'intensité de la douleur et la fatigue ont été inclus dans l'analyse quantitative.

Les résultats du critère de jugement « troubles du sommeil » n'ont pas pu être inclus dans l'analyse quantitative car l'échelle d'évaluation du symptôme est à l'envers par rapport aux autres articles inclus. Elle considère la valeur 0 comme très symptomatique et 100 comme asymptomatique.

- Heymann 2001 :

Seuls les résultats du critère de jugement « qualité de vie évaluée par la FIQ » sont présentés sous forme de moyennes avec écart type. Les résultats du critère de jugement « amélioration globale évaluée par le patient » sont présentés sous forme de pourcentage de patients améliorés et n'ont donc pas pu être inclus dans l'analyse quantitative.

Par ailleurs, il est à noter que certaines études ont rapporté les résultats des améliorations intragroupes pour l'amitriptyline et le placebo, et non la comparaison de leurs résultats finaux. Il s'agit de :

- Kempenaers 1994 : pour tous les résultats de l'étude dont douleur, fatigue, qualité du sommeil
- Don Goldenberg 1996 : (cf chapitre « résumé ») pour tous ses résultats dont amélioration globale par le patient et par le médecin, douleur, sommeil, fatigue, FIQ
- Ginsberg 1996 : pour les critères de jugements à l'exception de la fatigue soit : amélioration globale évaluée par le patient et le médecin, douleur et troubles du sommeil.

a) Evaluation de l'amélioration globale des symptômes par le patient

L'évaluation globale des symptômes de la fibromyalgie par le patient a été mesurée dans 11 études sur 12, selon deux approches complémentaires :

- Le « ressenti global du patient » mesuré principalement par EVA, mais aussi le VSIG et l'échelle d'amélioration discontinue (« pire », « inchangée », « amélioration minimale », « amélioration modérée », « amélioration marquée »)
- Les « répercussions globales de la pathologie sur la qualité de vie du patient » principalement mesurées par le FIQ

Il nous a paru pertinent de scinder ce critère de jugement selon ces deux parties.

2 études ont d'ailleurs mesuré l'amélioration globale du patient selon ces 2 approches :

- Don Goldenberg 1996 : amélioration globale mesurée par le patient par EVA et FIQ
- Heymann 2001 : amélioration globale mesurée par le patient par VSIG et FIQ

Plusieurs auteurs ont aussi utilisé des scores composites pour évaluer les répercussions globales de la pathologie sur la vie du patient. Bien que ces scores soient le résultat de plusieurs autoévaluations, nous avons décidé de ne pas les intégrer à notre critère de jugement car ils sont propres à chaque étude, décidés arbitrairement par les auteurs des essais, et donc sans garantie ni de fiabilité ni de reproductibilité. Nous n'avons pas non plus retenu les échelles de qualité de vie : Sickness Impact Profile SIP (Profile d'Impact de Maladie), Health Assessment Questionnaire HAQ (Questionnaire d'Evaluation de la santé), Sheehan's functional scales (Echelle Fonctionnelle de Sheehan), et Nottingham Health Profile dimension (Profile de Santé selon les Dimension de Nottingham) car ce ne sont pas des échelles spécifiques de la fibromyalgie.

a.1- Le ressenti global du patient

❖ Synthèse qualitative

L'analyse qualitative du « ressenti global du patient » est composée de 10 études : Carette 1986, Don Goldenberg 1986, Scudds 1989, Carette 1994, Kempnaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998 et Heymann 2001. Elle regroupe

approximativement 552 patients, avec environ 299 patients dans le groupe amitriptyline et 253 dans celui traité par placebo.

4 études rapportent des résultats montrant une différence significative en faveur des patients traités par amitriptyline par rapport à ceux traités par placebo : Don Goldenberg 1986 ($p < 0.05$), Scudds 1989 ($p < 0.001$), Carette 1995 ($p < 0.05$), Heymann 2001 ($p < 0.00981$).

3 études rapportent des résultats en faveur de l'amitriptyline par rapport au placebo, mais non significatifs : Carette 1986, Carette 1994, Hanonnen 1998.

3 études ne réalisent pas de comparaison entre les résultats finaux des groupes de traitement par amitriptyline et par placebo, mais mesurent l'amélioration intragroupe. Ces résultats montrent une amélioration significative dans le groupe amitriptyline : Kempenaers 1994 (amélioration intragroupe significative pour l'amitriptyline et pour le placebo ($p_{\text{time} \times \text{groupe}} < 0.05$)), Ginsberg 1996 (amélioration significative pour le groupe amitriptyline ($p_{\text{time}} < 0.001$) mais pas pour le groupe placebo (p non significatif)), Don Goldenberg 1996 (amélioration amitriptyline significative par rapport à l'amélioration dans le groupe placebo, p non rapportés dans l'article). Pour ces 3 études les résultats finaux sont en faveur de l'amitriptyline, mais pour les raisons évoquées ci-dessus, sans information de significativité.

Selon REB, aucun résultat n'est confirmatoire et moins de 50% des essais cliniques rapportent un résultat significatif pour ce critère de jugement.

❖ Synthèse quantitative

Toutes les études incluses comprennent des valeurs d'EVA finales évaluées sur une échelle de 0 à 10 cm ou sur une échelle de 0 à 100 mm. Aucun résultat n'est concluant pour ce critère de jugement.

L'analyse quantitative inclut 4 études et comprend 190 patients : 99 dans le groupe amitriptyline et 91 dans le groupe placebo. Les résultats finaux de ces études sont tous en faveur de l'amitriptyline, dont 1 de façon significative (Ginsberg 1996).

Le résultat de l'analyse quantitative est statistiquement significatif ($p = 0.003$). La standardized mean difference (SMD) est à -0.68 avec un intervalle de confiance à 95% compris entre [-1.13 ; -0.24]. L'hétérogénéité est de $I^2 = 55\%$ avec $p = 0.08$.

La SMD est donc significativement en faveur de l'amitriptyline par rapport au placebo avec une taille d'effet modérée pour ce critère de jugement.

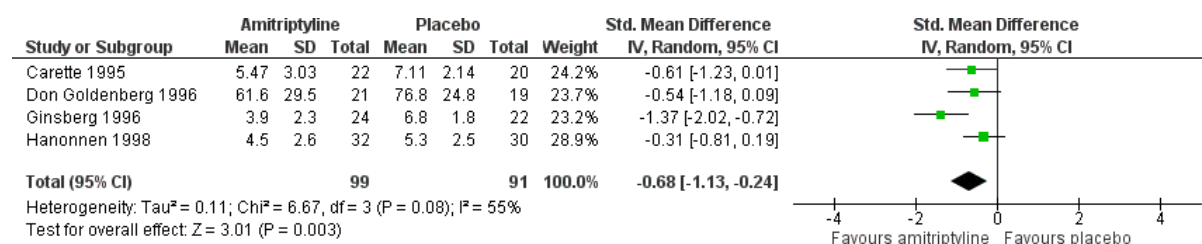


Figure 4 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : Amélioration globale évaluée par le patient – ressenti du patient

a.2- Les répercussions de la pathologie sur la qualité de vie du patient

❖ Synthèse qualitative

L'analyse qualitative est composée de 3 études mesurant l'amélioration globale des symptômes évaluée par le patient à travers les « répercussions globales de la pathologie sur la qualité de vie du patient » : Don Goldenberg 1996, Heymann 2001 et Braz 2012. Elle regroupe 136 patients, avec 71 patients dans le groupe amitriptyline et 65 dans celui traité par placebo.

1 étude rapporte des résultats en faveur de l'amitriptyline mais non significatifs : Heymann 2001. Bien que ce critère de jugement soit le principal de cet essai clinique, étant donné que le résultat n'est pas significatif, il n'est pas confirmatoire.

1 étude (Don Goldenberg 1996) ne réalise pas de comparaison entre les résultats finaux des patients traités par amitriptyline et ceux traités par placebo. Le résultat pour ce critère de jugement montre une amélioration statistiquement significative dans le groupe amitriptyline par rapport à celle du groupe placebo. Pour cet essai clinique, les résultats finaux de ce critère de jugement sont en faveur de l'amitriptyline, mais pour les raisons évoquées ci-dessus, sans information de significativité.

1 étude ne montre aucune différence entre le traitement par amitriptyline et placebo : Braz 2012.

Selon REB, aucun résultat n'est confirmatoire, et moins de 50% sont significatifs pour ce critère de jugement

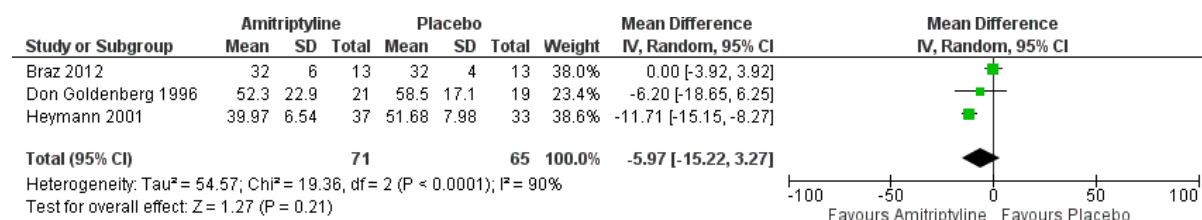
❖ Synthèse quantitative

Le FIQ est composé de 10 items avec une note maximale à 10 pour chacun des items. La note maximale du FIQ est donc 100. Plus la note est élevée plus l'impact sur la personne est important. Aucun résultat n'est concluant pour ce critère de jugement.

L'analyse quantitative inclut 3 études et comprend 136 patients : 71 dans le groupe amitriptyline et 65 dans le groupe placebo. Parmi les résultats finaux de ces études, 2 sont en faveur de l'amitriptyline, dont 1 de façon significatif (Heymann 2001). Les résultats finaux de l'étude Braz 2012 ne montrent pas de différence entre le traitement par amitriptyline et par placebo pour ce critère de jugement.

Le résultat de l'analyse quantitative n'est pas statistiquement significatif ($p = 0.21$). La SMD est à -5,97 avec un intervalle de confiance à 95% compris entre [-15.22 ; 3.27]. L'hétérogénéité est de $I^2 = 90\%$ avec $p < 0.0001$.

Le résultat est en faveur de l'amitriptyline par rapport au placebo mais il n'est pas statistiquement significatif pour ce critère de jugement.



Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : Amélioration globale évaluée par le patient – répercussions sur la qualité de vie

b) Evaluation de l'amélioration globale par le médecin

❖ Synthèse qualitative

L'évaluation globale des symptômes de la fibromyalgie par le médecin a été mesurée dans 8 études, principalement par l'EVA, mais aussi par le « Clinical impression of change », l'échelle « Clinical Impression of Severity » et l'échelle d'amélioration (« pire », « inchangée », « amélioration minimale », « amélioration modérée », « amélioration marquée »). Elle regroupe approximativement 405 patients, avec environ 224 patients dans le groupe amitriptyline et 181 dans celui traité par placebo.

1 étude (Hanonnen 1998) a utilisé 2 échelles de mesure différentes pour ce critère de jugement :

- L'une en critère de jugement principal (Clinical impression of Change) qui montre une amélioration significative du groupe amitriptyline par rapport au placebo de l'échelle de mesure ($p = 0.003$) et du nombre de patients améliorés ($p = 0.017$)
- L'autre en critère de jugement secondaire : (Clinical Impression of Severity par le médecin) qui montre une amélioration non significative en faveur de l'amitriptyline.

Le résultat de l'étude est donc en faveur d'une amélioration significative du critère de jugement « amélioration des symptômes mesurée par le médecin ». Cependant pour l'analyse quantitative nous avons fait le choix d'inclure les résultats du critère de jugement secondaire pour plus d'homogénéité dans les échelles de mesures.

3 études rapportent des résultats montrant une amélioration significative des patients traités par amitriptyline par rapport à ceux traités par placebo : Don Goldenberg 1986 ($p < 0.01$), Carette 1995 ($p < 0.05$), Hanonnen 1998 ($p = 0.017$).

2 études rapportent des résultats en faveur de l'amitriptyline par rapport au placebo, mais non significatif ou sans information sur la significativité du résultat : Carette 1986 (pas d'information sur p), Carette 1994 (p non significatif).

2 études ne réalisent pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo, mais mesurent l'amélioration intragroupe et trouvent une amélioration significative dans le groupe amitriptyline : Kempenaers 1994 (amélioration intragroupe significative pour l'amitriptyline et pour le placebo ($p \text{ time } \times \text{ groupe } < 0.05$)),

Ginsberg 1996 (amélioration significative pour le groupe amitriptyline (p time < 0.001) mais pas pour le groupe placebo (p non significatif)). Pour ces 2 études les résultats finaux sont en faveur de l'amitriptyline, mais pour les raisons évoquées ci-dessus, sans information de significativité.

1 étude (Don Goldenberg 1996) ne réalise pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo. Elle rapporte dans ses résultats une amélioration intragroupe significative pour l'amitriptyline ($p = 0.04$), mais en comparaison avec l'amélioration dans le groupe placebo la différence n'est pas significative (p non rapporté). Etant donné que cette différence n'est pas significative, et bien que les valeurs finales soient en faveur de l'amitriptyline, l'étude est considérée comme ne rapportant pas de différence entre les 2 traitements pour ce critère.

Parmi ces études, seule l'étude de Hanonnen 1998 est considérée comme étant confirmatoire pour la mesure de ce critère de jugement mesuré par l'échelle « Clinical Impression of Changment ». Cependant il est important de prendre en compte le fait que cet essai est à haut risque de biais et n'est donc pas concluant. De plus, il étudie une autre molécule en priorité (Nortriptyline) et l'amitriptyline est considérée comme le traitement de référence. A noter que pour l'analyse quantitative ci-dessous, nous n'avons pas pu utiliser ce résultat et avons eu recours à celui du critère de jugement secondaire Clinical impression of Severity.

Selon REB, l'étude Hanonnen 1998 rapporte un résultat confirmatoire mais non concluant, et moins de 50% des résultats pour ce critères de jugement sont significatifs.

❖ Synthèse quantitative

Toutes les études incluses comprennent des valeurs d'EVA finales évaluées sur une échelle de 0 à 10 cm ou sur une échelle de 0 à 100 mm ou des valeurs de l'échelle « Impression Clinique de Sévérité » mesurées par le médecin de 0 à 7. Aucun résultat n'est concluant pour ce critère de jugement.

La méta-analyse regroupe 4 études et comprend 188 patients : 98 dans le groupe amitriptyline et 90 dans le groupe placebo. Parmi les résultats finaux de ces études, 4 sont en faveur de l'amitriptyline, dont 1 de façon significatif (Ginsberg 1996).

Le résultat de l'analyse quantitative est statistiquement significatif ($p < 0.0001$). La SMD est à -0.63 avec un intervalle de confiance à 95% compris entre [-0.93 ; -0.34]. L'hétérogénéité est de $I^2 = 38\%$ avec $p = 0.18$.

La SMD est donc significativement en faveur de l'amitriptyline par rapport au placebo avec une taille d'effet « modéré » pour ce critère de jugement.

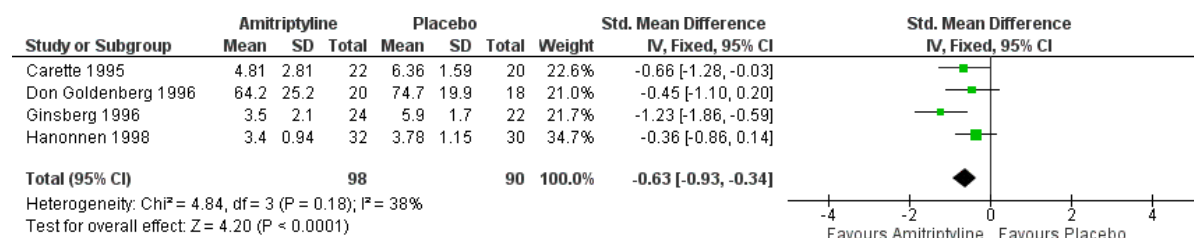


Figure 5 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : Amélioration globale évaluée par le médecin

c) Evaluation de la douleur

❖ Synthèse qualitative

L'intensité de la douleur a été mesurée dans 11 études, principalement par EVA, mais aussi par le score de Mc Gill. Elle regroupe approximativement 557 patients, avec environ 297 patients dans le groupe amitriptyline et 260 dans celui traité par placebo.

3 études rapportent des résultats montrant une amélioration significative des patients traités par amitriptyline par rapport à ceux traités par placebo : Don Goldenberg 1986 ($p < 0.01$), Scudds ($p < 0.05$), Carette 1995 ($p < 0.05$)

4 études rapportent des résultats en faveur de l'amitriptyline par rapport au placebo, mais non significatif : Carette 1986, Carette 1994, Hanonnen 1998, Fors 2002.

2 études ne réalisent pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo, mais mesurent l'amélioration intragroupe et trouvent une amélioration significative dans le groupe amitriptyline : Ginsberg 1996 (amélioration significative groupe amitriptyline (p time < 0.001) mais pas dans le groupe placebo (p non significatif)), ainsi que Don Goldenberg 1996 (amélioration significative en intragroupe amitriptyline ($p = 0.02$), et comparée à l'amélioration du placebo, p non rapporté). Pour ces 2

études les résultats finaux sont en faveur de l'amitriptyline, mais pour les raisons évoquées ci-dessus, sans information de significativité.

1 étude ne réalise pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo mais rapporte une amélioration intragroupe non significative pour l'amitriptyline et pour le placebo : Kempenaers 1994. Etant donné que cette différence n'est pas significative, et bien que les valeurs finales soient en faveur de l'amitriptyline, l'étude est considérée comme ne rapportant pas de différence entre les 2 traitements pour ce critère.

1 étude rapporte un résultat en faveur du placebo par rapport à l'amitriptyline, mais non significatif : Braz 2012.

1 étude (Carette 1994) a utilisé 2 échelles de mesure différentes (EVA et score de Mc Gill) pour ce critère de jugement, toutes 2 en critère de jugement secondaire et concordantes dans leurs résultats : amélioration non significative des patients traités par amitriptyline par rapport au groupe placebo.

Pour l'analyse quantitative, étant donné que les résultats mesurés par EVA sont des estimations de moyennes et que les écarts types n'ont pu être ni calculés ni estimés, nous avons fait le choix d'inclure les résultats de « intensité de la douleur au moment présent » du score de Mc Gill pour l'analyse quantitative.

Selon REB, les résultats de l'intensité de la douleur ne sont considérés comme confirmatoires dans aucune étude de notre analyse qualitative. En effet, l'étude de Scudds 1989 a bien pour critère de jugement principal un score composite évaluant la douleur et dont la mesure de l'intensité est une des composantes. De plus les auteurs ont utilisé la méthode de Tukey pour ajuster le risque alpha. Le résultat du score composite est donc confirmatoire. Mais les scores qui le composent sont analysés indépendamment en critères de jugement secondaires, et sont donc exploratoires. Par ailleurs, moins de 50% des études incluses rapportent une différence statistiquement significative lors de la comparaison des résultats finaux des patients traités par amitriptyline et par placebo pour ce critère de jugement.

❖ Synthèse quantitative

Les études incluses comprennent des valeurs d'EVA finales évaluées sur une échelle de 0 à 10 cm ou sur une échelle de 0 à 100 mm et l'estimation de l'intensité par le score de Mc Gill à l'instant du questionnaire sur une échelle discontinue de 1 à 5. Aucun résultat n'est concluant pour ce critère de jugement.

L'analyse quantitative inclut 9 études et comprend 452 patients : 245 dans le groupe amitriptyline et 207 dans le groupe placebo. Parmi les résultats finaux de ces études, 7 sont en faveur de l'amitriptyline, dont 3 de façon significative (Carette 1995, Don Goldenberg 1996 et Ginsberg 1996). L'étude de Kempnaers ne semble pas être en faveur d'une différence entre le traitement par amitriptyline et placebo pour ce critère de jugement. Les résultats finaux d'1 étude sont en faveur du placebo et sont statistiquement significatifs : Braz 2012.

Le résultat de l'analyse quantitative est statistiquement significatif ($p = 0.03$). La SMD est à -0.42 avec un intervalle de confiance à 95% compris entre [-0.79 ; -0.04]. L'hétérogénéité est de $I^2 = 71\%$ avec $p = 0.0006$.

La SMD est donc significativement en faveur de l'amitriptyline par rapport au placebo avec une taille d'effet « faible » à « modéré ».

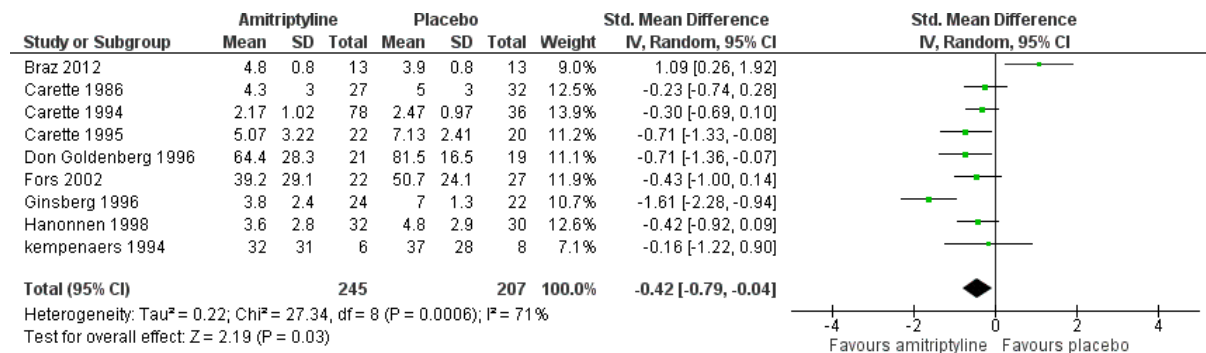


Figure 6 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : Intensité de la douleur

d) Evaluation du sommeil

❖ Synthèse qualitative

Le sommeil a été mesuré dans 9 études, principalement par EVA, mais aussi par une échelle d'amélioration (« pire », « inchangé », « mieux ». Elle regroupe approximativement 433 patients, avec environ 238 patients dans le groupe amitriptyline et 195 dans celui traité par placebo.

4 études rapportent des résultats montrant une amélioration significative des patients traités par amitriptyline par rapport à ceux traités par placebo : Carette 1986 ($p = 0.02$), Don Goldenberg 1986 ($p < 0.05$), Carette 1995 ($p < 0.05$), Hanonnen 1998 ($p < 0.01$)

1 étude rapporte des résultats en faveur de l'amitriptyline par rapport au placebo, mais non significatif : Carette 1994

2 études ne réalisent pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo, mais mesurent l'amélioration intragroupe et trouvent une amélioration significative dans le groupe amitriptyline : Ginsberg 1996 (amélioration significative groupe amitriptyline ($p_{\text{time}} < 0.001$) mais pas groupe placebo (p non significatif)) et Don Goldenberg 1996 (amélioration significative groupe amitriptyline ($p < 0.001$), et significative comparée au placebo (p non rapporté)). Pour ces 2 études les résultats finaux sont en faveur de l'amitriptyline, mais pour les raisons évoquées ci-dessus, sans information de significativité.

1 étude ne réalise pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo mais rapporte une amélioration intragroupe non significative pour l'amitriptyline et pour le placebo : Kempenaers 1994. Etant donné que cette différence n'est pas significative, bien que les valeurs finales soient en faveur du placebo, l'étude est considérée comme ne rapportant pas de différence entre les 2 traitements pour ce critère.

1 étude rapporte un résultat en faveur du placebo par rapport à l'amitriptyline, mais non significatif : Braz 2012

Selon REB, aucun résultat n'est confirmatoire, et moins de 50% sont significatifs pour ce critère de jugement

❖ Synthèse quantitative

Toutes les études incluses comprennent des valeurs d'EVA finales évaluées sur une échelle de 0 à 10 cm ou sur une échelle de 0 à 100 mm. Aucun résultat n'est concluant pour ce critère de jugement.

L'analyse quantitative inclut 5 études et comprend 216 patients : 112 dans le groupe amitriptyline et 104 dans le groupe placebo. Parmi les résultats finaux de ces études, 4 sont en faveur de l'amitriptyline, dont 2 de façon significative (Carette 1995 et Ginsberg 1996). Les résultats de l'étude Braz 2012 sont en faveur du placebo mais ne sont pas statistiquement significatifs.

Le résultat de l'analyse qualitative est statistiquement significatif ($p = 0.0002$). La SMD est à -0.53 avec un intervalle de confiance à 95% compris entre [-0.80 ; -0.25]. L'hétérogénéité est de $I^2 = 31\%$ avec $p = 0.22$.

La SMD est donc significativement en faveur de l'amitriptyline par rapport au placebo avec une taille d'effet « modéré ».

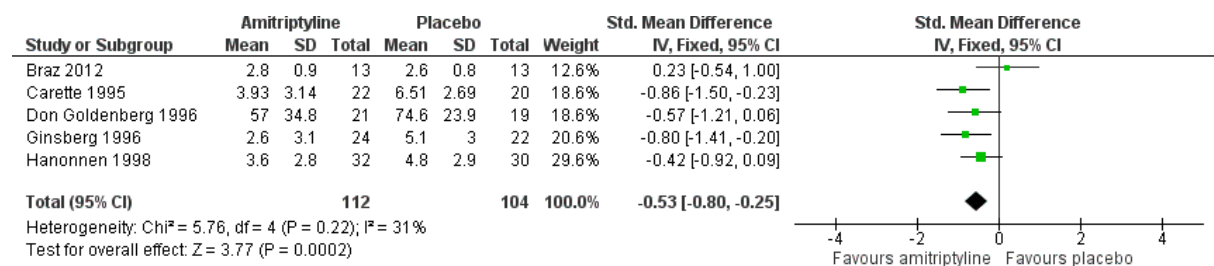


Figure 7 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : Sommeil

e) Evaluation de la fatigue

❖ Synthèse qualitative

La fatigue a été mesurée dans 8 études, principalement par EVA, mais aussi par une échelle d'amélioration (« pire », « inchangée », « mieux »). Elle regroupe approximativement 374 patients, avec environ 211 patients dans le groupe amitriptyline et 163 dans celui traité par placebo.

3 études rapportent des résultats montrant une amélioration significative des patients traités par amitriptyline par rapport à ceux traités par placebo : Don Goldenberg 1986 ($p < 0.05$), Carette 1995 ($p < 0.05$), Ginsberg 1996 ($p < 0.014$).

1 étude rapporte des résultats en faveur de l'amitriptyline par rapport au placebo, mais non significatif : Carette 1994.

1 étude ne réalise pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo, mais mesure l'amélioration intragroupe et trouve une amélioration significative dans le groupe amitriptyline : Kempenaers 1994 (amélioration intragroupe significative pour l'amitriptyline et pour le placebo ($p \text{ time } \times \text{ groupe } < 0.05$)). Pour cette étude les résultats finaux sont en faveur de l'amitriptyline, mais pour les raisons évoquées ci-dessus, sans information de significativité.

1 étude ne réalise pas de comparaison entre les résultats finaux des patients traités par amitriptyline et par placebo, mais rapporte une amélioration intragroupe non significative pour l'amitriptyline ($p = 0.26$), et non significative lors de la comparaison au groupe placebo (p non rapporté) : Don Goldenberg 1996. Etant donné que cette différence n'est pas significative, et bien que les valeurs finales soient en faveur de l'amitriptyline, l'étude est considérée comme ne rapportant pas de différence entre les 2 traitements pour ce critère.

2 études rapportent un résultat en faveur du placebo par rapport à l'amitriptyline, mais non significatif : Hanonnen 1998 et Braz 2012

Selon REB, aucun résultat n'est confirmatoire, et moins de 50% sont significatifs pour ce critère de jugement

❖ Synthèse quantitative

Toutes les études incluses comprennent soit des valeurs d'EVA finales évaluées sur une échelle de 0 à 10 cm ou sur une échelle de 0 à 100 mm soit une évaluation par une échelle de 0 à 3. Aucun résultat n'est concluant pour ce critère de jugement.

La méta-analyse inclus 6 études et comprend 230 patients : 118 dans le groupe amitriptyline et 112 dans le groupe placebo. Parmi les résultats finaux de ces études, 3 sont en faveur de l'amitriptyline, dont 2 de façon significatif (Carette 1995 et Ginsberg 1996). Les études de Don

Goldenberg 1996 et Hanonnen 1998 ne semble pas être en faveur d'une différence entre le traitement par amitriptyline et placebo pour ce critère de jugement Les résultats de l'étude Braz 2012 sont en faveur du placebo mais ne sont pas statistiquement significatifs.

Le résultat de l'analyse quantitative n'est pas statistiquement significatif ($p = 0.13$). La SMD est à -0.36 avec un intervalle de confiance à 95% compris entre [-0.82 ; 0.11]. L'hétérogénéité est de $I^2 = 65\%$ avec $p = 0.01$.

Le résultat est en faveur de l'amitriptyline par rapport au placebo mais il n'est pas statistiquement significatif.

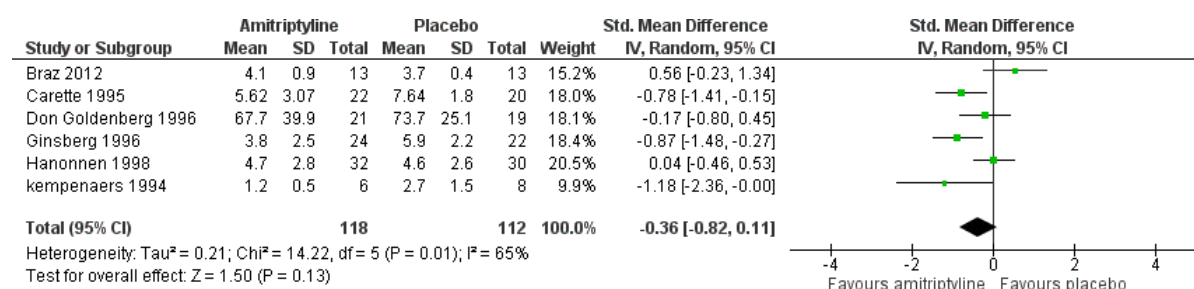


Figure 8 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : Fatigue

f) Evaluation des effets secondaires et des perdus de vue

a.1- Les effets indésirables

❖ Synthèse qualitative

Les effets indésirables ont été mesurés dans 7 études. L'étude Don Goldenberg 1996 mentionne uniquement avoir constaté des différences en termes de type d'effets indésirables entre les 4 groupes de traitements. Etant donné que ni les résultats ni la significativité ne sont rapportés dans l'article, nous avons décidé de ne pas l'inclure ni dans l'analyse qualitative ni dans l'analyse quantitative. L'analyse qualitative porte donc sur 6 études et regroupe approximativement 444 patients, avec environ 241 patients dans le groupe amitriptyline et 203 dans celui traité par placebo.

1 étude rapporte des résultats montrant une augmentation significative des effets indésirables parmi les patients traités par amitriptyline par rapport à ceux traités par placebo : Ginsberg 1996

2 études rapportent des résultats non significatifs ou sans information sur leur significativité, montrant une augmentation des effets indésirables parmi les patients traités par amitriptyline par rapport à ceux traités par placebo : Carette 1986, Carette 1994.

2 études rapportent des résultats montrant une absence de différence en termes de fréquence entre les effets indésirables par rapport au placebo : Don Goldenberg 1986, Heymann 2001

1 étude rapporte un nombre d'effets indésirables moins important dans le groupe amitriptyline que dans le groupe placebo : Hanonnen 1998.

Par ailleurs, la nature des effets indésirables diffère entre les patients traités par amitriptyline (bouche sèche, somnolence) et le placebo dans 4 études : Carette 1986, Don Goldenberg 1986 (placebo et naproxène + placebo), Ginsberg 1996, Hanonnen 1998.

Les auteurs ne retrouvent pas clairement de différence en termes de nature des effets indésirables en considérant bouche sèche et somnolence entre les patients traités par placebo ou par l'amitriptyline pour 1 étude : Heymann 2001.

L'étude Carette 1994 ne donne pas d'information sur la nature des effets indésirables entre l'amitriptyline et le placebo.

❖ Synthèse quantitative

Bien que presque toutes les études évoquent les effets secondaires, seules trois présentent des données détaillées sur les effectifs de patients ayant présenté des effets secondaires

La méta-analyse comprend 185 patients : 91 dans le groupe amitriptyline et 94 dans le groupe placebo. Le risque ratio est à 2.93 avec un intervalle de confiance à 95% compris entre [0.37 ; 23.42]. L'hétérogénéité est de $I^2 = 92\%$ avec $p < 0.00001$.

Le résultat est en faveur d'une augmentation du nombre d'effets indésirables dans le groupe amitriptyline par rapport au groupe placebo, mais il n'est pas statistiquement significatif.

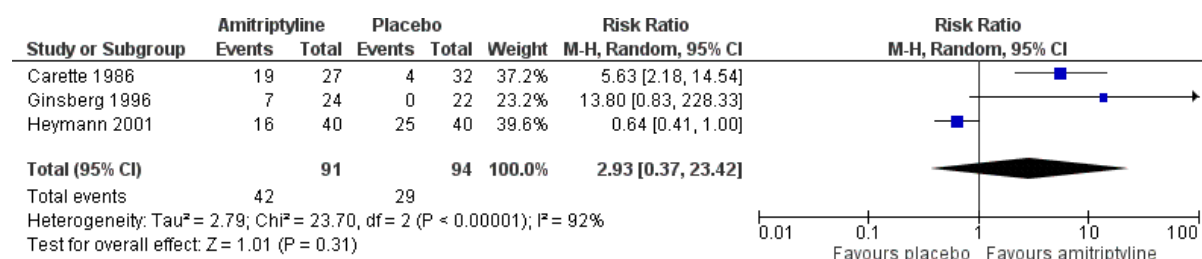


Figure 9 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : effets indésirables

a.2- Les perdus de vue

❖ Synthèse quantitative

Toutes les études incluses dans l'analyse qualitative rapportent leurs nombres de perdus de vue par groupe de traitement, à l'exception de Don Goldenberg 1986 et Kempenaers 1994.

Etant donné que les informations sur les perdus de vue ne sont pas un critère de jugement, mais un critère de qualité méthodologique, la notion de résultat exploratoire ou confirmatoire n'a

pas de sens dans ce cas. En cela, une analyse qualitative n'est pas pertinente. 2 études ont fait l'objet d'une discussion entre les auteurs pour arriver à un consensus :

- Carette 1995 : essai en cross-over. Les 2 perdus de vue sont survenus pendant la période de wash out. Nous avons décidé de n'attribuer ces perdus de vue à aucun des 2 bras de traitements.
- Ginsberg 1996 : les auteurs rapportent avoir perdu de vue 5 patients entre la première visite permettant l'inclusion des patients et la première visite permettant la mesure des résultats intermédiaires. Ils ont considéré que ces patients ne faisaient pas partie de l'étude. Etant donné que les patients avaient été randomisés et avaient reçu les médicaments, nous avons décidé de les considérer comme des perdus de vue, au même titre que les 4 perdus de vue survenus dans la suite de l'étude. Les patients ayant eu une mauvaise observance n'ont pas été considérés comme perdus de vue.

L'analyse quantitative regroupe donc 10 études sur 12, et rapporte le nombre de patients perdus de vue dans les groupes de traitements amitriptyline et placebo, par rapport à leurs effectifs respectifs à la randomisation. Elle comprend 663 patients (effectif à la randomisation) avec 351 dans le groupe amitriptyline et 312 dans le groupe placebo. Toutes ces études sont considérées comme étant exploratoires pour ce critère de jugement.

Le résultat est non significatif ($p = 0.05$). La SMD est à 0.69 avec un intervalle de confiance à 95% compris entre [0.48 ; 1]. L'hétérogénéité est de $I^2 = 0\%$ avec $p = 0.72$.

Le résultat est donc non significatif, mais suggère une tendance à une augmentation du nombre de perdus de vue dans le groupe placebo par rapport au groupe amitriptyline, avec une taille d'effet « modérée ».

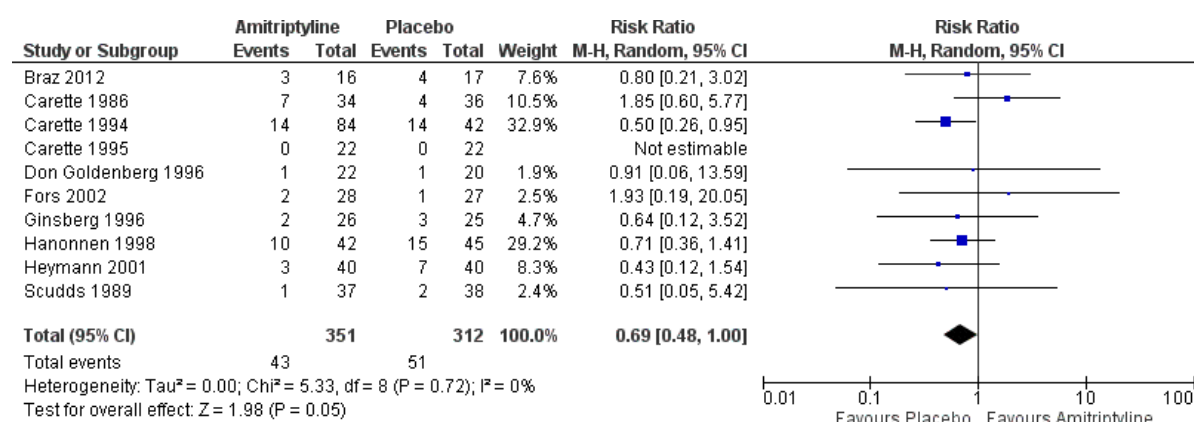


Figure 10 : Forest plot de comparaison : Amitriptyline versus Placebo, critère de jugement : perdus de vue

6. Evaluation du niveau de preuve global selon REB et GRADE

L'évaluation du niveau de preuve des analyses quantitatives ci-dessus a été réalisée par 2 méthodes : le GRADE (15) qui est la méthode de référence utilisée dans toutes les méta-analyses Cochrane, et le REB (14), une nouvelle méthode plus rigoureuse car elle tient compte des domaines de biais du GRADE mais aussi du risque de biais lié à l'inflation du risque alpha.

Pour la méthode GRADE, le niveau de preuve dépend de 5 domaines de qualité qui sont d'égale importance. La méthode estime initialement tout essai clinique comme ayant un niveau de preuve maximal (soit élevé). Les auteurs de l'étude quantitative retirent ensuite pour chacun des 5 domaines de qualité entre 0, 1 ou 2 points selon le risque estimé. L'analyse quantitative perd des niveaux de preuve jusqu'à être estimée au pire comme étant de « très faible » niveau de preuve (ce qui correspond à la perte de 3 points ou plus). Or tous nos essais cliniques ont un niveau de risque de biais élevé, ce qui nous a fait retirer systématiquement 2 points pour ce domaine. De plus, l'analyse des biais de publication fait intervenir l'évaluation du funnel plot de chaque analyse quantitative. Mme Veloso nous a aidé à interpréter ces figures. Au vu du faible nombre d'études incluses dans chaque analyse quantitative, il est difficile de juger de la symétrie ou de l'asymétrie de chaque funnel plot, ce qui rend leur interprétation limitée. Ainsi, nous avons décidé d'estimer qu'il y a systématiquement un risque de biais de publication pour chaque analyse quantitative réalisée, et avons retiré 1 point sur le niveau de preuve selon le GRADE. Toutes nos analyses quantitatives sont donc estimées comme étant de très faible niveau de preuve selon le GRADE. Le détail du niveau de preuve pour chaque analyse quantitative est présenté ci-dessous.

Pour la méthodologie REB, le niveau de preuve d'une analyse quantitative dépend prioritairement du niveau de preuve de chaque résultat qui la compose, c'est-à-dire du nombre de résultats conclusifs et du pourcentage de résultats statistiquement significatifs inclus. Un résultat conclusif est un confirmatoire obtenu dans un essai clinique dont le risque de biais global estimé est faible. Aucun de nos résultats n'est conclusif. En l'absence de résultat conclusif, si moins de la moitié des résultats sont significatifs, le niveau de preuve de l'analyse qualitative est estimé « manque de preuve ». C'est le cas pour tous nos critères de jugement.

Le détail du niveau de preuve selon ces 2 méthodologies est présenté ci-dessous, pour les analyses quantitatives : amélioration globale évaluée par le patient, par le médecin, douleur,

troubles du sommeil et fatigue. Il ne nous a pas semblé pertinent de vérifier le niveau de preuve pour les analyses quantitatives des effets secondaires et des perdus de vue.

a) Evaluation de l'amélioration globale des symptômes par le patient

a.1- Ressenti du patient

❖ Selon REB

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2. Aucune étude n'est concluante pour ce critère de jugement. Moins de la moitié des essais cliniques analysés montrent un résultat statistiquement significatif. Notre méta-analyse de sensibilité, retrouve une différence de moyenne standardisée significative à -0.68 en faveur de l'amitriptyline mais avec une hétérogénéité élevée et un test I^2 à 55%. Cependant notre analyse qualitative inclut seulement 4 études parmi les 10 identifiées lors de l'analyse qualitative. De plus le faible nombre d'études incluses rend l'interprétation du *funnel plot* non fiable, bien que ce dernier semble être plutôt symétrique. Ces éléments ne nous permettent pas d'affirmer qu'il n'existe pas de biais de publication.

Selon la méthode REB, on considère qu'il y a un “manque de preuve” quant à l'efficacité de l'amitriptyline pour le critère de jugement « amélioration globale - ressenti du patient ».

❖ Selon GRADE

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

Risque de biais : Les quatre études incluses dans notre méta-analyse présentent un risque élevé de biais selon l'outil RoB 2. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, deux études Carrette 1995 et Ginsberg 1996 présentent un risque jugé élevé pour 3 domaines du RoB 2 et les deux autres Don Goldenberg 1996 et de Hanonnen 1998 ont un risque de biais élevé pour deux domaines. De ce fait nous décidons de rétrograder le niveau de preuve de deux niveaux.

Incohérence : Les 4 études suggèrent toutes une différence en faveur de l'amitriptyline. Cependant, le coefficient d'hétérogénéité I² est à 55% avec un coefficient de significativité $p = 0.08$. L'hétérogénéité est donc modérée, mais non significative. Nous décidons de ce fait de retirer 1 point au niveau de preuve

Caractère indirect des résultats : Les critères de jugement des quatre études et leurs populations sont en lien avec l'objectif de notre revue de la littérature. Parmi les 4 études, 3 études mesurent l'efficacité d'une autre molécule en même temps que l'amitriptyline par rapport au placebo. Parmi les 4 études, 2 études ne comparent pas les résultats finaux des patients traités par amitriptyline et par placebo mais observent l'amélioration intragroupe. Nous décidons de ce fait de retirer 1 point.

Imprécision : L'effectif total de notre méta-analyse est de 190 patients et est donc considéré comme faible. Notre SMD est -0.68 avec un intervalle de confiance compris entre [-1.13 ; -0.24]. Elle est donc à moins de 0.5 des limites de l'intervalle de confiance et ses bornes sont cohérentes pour situer l'effet de l'amitriptyline comme étant favorable par rapport au placebo à 95%. De ce fait nous considérons qu'il existe un risque sérieux et nous décidons d'enlever 1 point au niveau de preuve, à cause de l'effectif

Biais de publication : Les études incluses sont toutes à petits échantillons. Le *funnel plot* ne rassemble pas suffisamment d'études pour permettre une interprétation fiable, bien qu'il semble plutôt montrer une distribution symétrique des études. De ce fait, nous ne pouvons exclure un biais de publication. Nous enlevons 1 points au niveau de preuve.

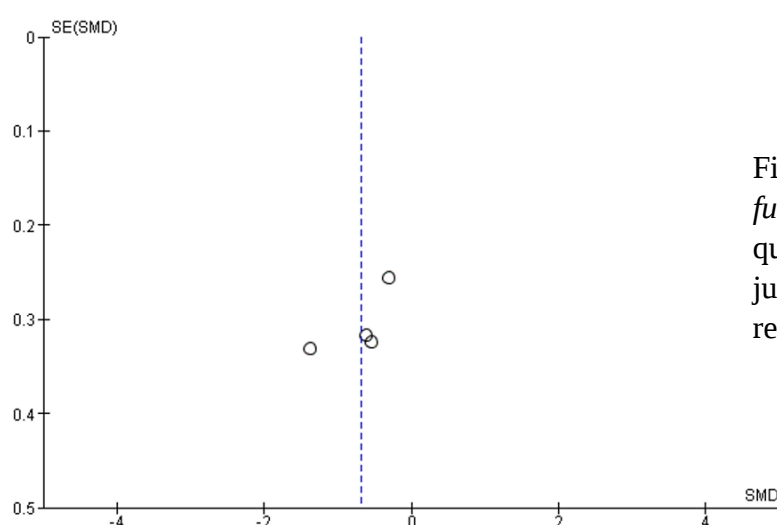


Figure 11 : Représentation du *funnel plot* de l'analyse quantitative pour le critère de jugement « amélioration globale - ressenti du patient »

Nous n'avons pas rajouté de points pour les items « ampleur de l'effet », « effet dose » et « biais plausibles ».

Conclusion : En prenant en compte ces résultats, nous concluons que l'efficacité de l'amitriptyline est de « très faible » niveau de preuve pour l'amélioration globale des patients selon leur ressenti.

a.2- Répercussion de la pathologie sur la qualité de vie

❖ Selon REB

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2. Aucune étude n'est concluante pour ce critère de jugement. Moins de la moitié des essais cliniques analysés montrent un résultat statistiquement significatif. Notre méta-analyse de sensibilité, retrouve une différence de moyennes standardisées à - 5,97 en faveur de l'amitriptyline mais le résultat n'est pas statistiquement significatif. Par ailleurs, aucune étude complète non publiée n'a été identifiée sur le site clinicaltrials.gov, mais le faible nombre d'études incluses rend l'interprétation du *funnel plot* non fiable, bien que ce dernier semble être plutôt symétrique. Ces éléments ne nous permettent pas d'affirmer qu'il n'existe pas de biais de publication. Selon la méthode REB, on considère qu'il y a un “ manque de preuve” quant à l'efficacité de l'amitriptyline pour le critère de jugement « l'amélioration globale – répercussion de la pathologie sur la qualité de vie du patient ».

❖ Selon GRADE

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

Risque de biais : Les trois études incluses dans notre méta-analyse présentent un risque élevé de biais selon l'outil RoB 2. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, L'étude de Braz 2012 présente un risque jugé élevé pour 3 domaines du RoB 2. Pour Don Goldenberg 1996 le risque de biais est élevé pour deux domaines et pour Heymann 2001 un domaine est à risque élevé. De ce fait nous décidons de rétrograder le niveau de preuve de deux niveaux.

Incohérence : 2 études suggèrent une différence en faveur de l'amitriptyline : Heymann 2001 et Don Goldenberg 1996. L'étude de Braz 2012 suggère une absence de différence entre le

traitement par amitriptyline et par placebo. Ces incohérences expliquent une très forte hétérogénéité pour la méta-analyse avec un test I^2 à 90% avec un $p < 0.0001$. Nous décidons de ce fait qu'il existe un risque sérieux et nous décidons de rétrograder le niveau preuve de 2 niveaux.

Caractère indirect des résultats : Les critères de jugement des 3 études et leurs populations sont en lien avec l'objectif de notre revue de la littérature. Les 3 études comparent chacune une autre molécule active en plus de l'amitriptyline au placebo. 1 étude ne compare pas les résultats finaux des patients traités par amitriptyline et par placebo mais observe l'amélioration intragroupe. Nous décidons de ce fait de retirer 1 point au niveau de preuve.

Imprécision : L'effectif total de notre méta-analyse est de 136 patients et est donc considéré comme faible. La différence de moyennes standardisées est non significative ($p = 0.21$) de -5,97 avec un intervalle de confiance compris entre [-15.22 ; 3.27]. L'intervalle de confiance montre une imprécision dans la mesure avec des bornes à plus de 0.5 de la taille d'effet. De plus, il est incohérent car il contient des valeurs en faveur et en défaveur de l'amitriptyline. De ce fait nous considérons qu'il existe un risque sérieux et nous enlevons 2 points niveau preuve.

Biais de publication : Les études sont toutes à petits échantillons. Le *funnel plot* ne rassemble pas suffisamment d'études pour permettre une interprétation fiable, bien qu'il semble plutôt montrer une distribution symétrique des études. De ce fait, et bien que la recherche d'études complètes sur la plateforme clinicaltrials.gov se soit avérée négative, nous ne pouvons exclure un biais de publication. Nous enlevons 1 point au niveau de preuve.

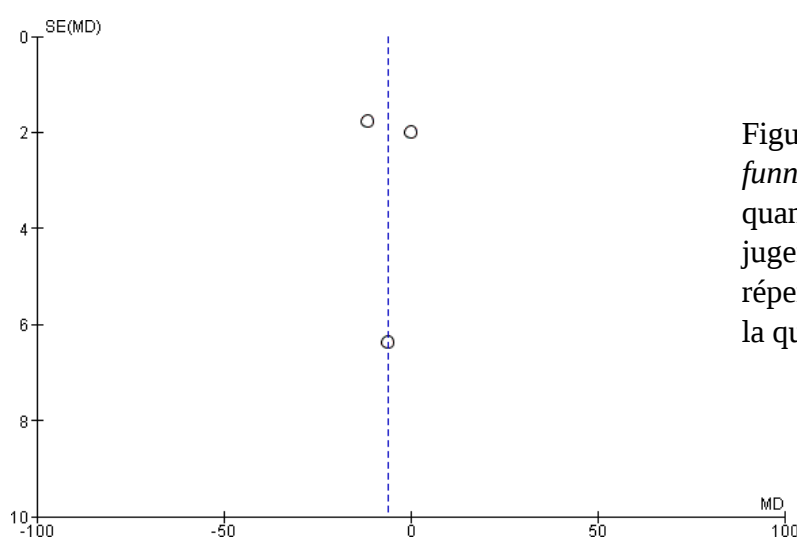


Figure 12 : Représentation du *funnel plot* de l'analyse quantitative pour le critère de jugement « amélioration globale – répercussions de la pathologie sur la qualité de vie »

Nous n'avons pas rajouté de points pour les items “ ampleur de l'effet”, “ effet dose” et “ biais plausibles “.

Conclusion : En prenant en compte ces résultats, nous concluons que l'efficacité de l'amitriptyline est de « très faible » niveau de preuve pour l'amélioration globale des patients – répercussion de la pathologie sur la qualité de vie évaluée par le patient à l'aide du FIQ.

b) Evaluation de l'amélioration globale par le médecin

❖ Selon REB

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2. Aucune étude n'est concluante pour ce critère de jugement. Moins de la moitié des essais cliniques analysés montrent un résultat statistiquement significatif.

De plus, notre méta-analyse de sensibilité retrouve une différence de moyennes standardisées significative à -0.63 en faveur de l'amitriptyline. L'hétérogénéité est modérée avec un test I^2 à 38%. Cependant notre analyse qualitative inclut seulement 4 études parmi les 8 identifiées lors de l'analyse qualitative. De plus, le faible nombre d'études incluses rend l'interprétation du *funnel plot* non fiable, bien que ce dernier semble être plutôt symétrique. Ces éléments ne nous permettent pas d'affirmer qu'il n'existe pas de biais de publication.

Selon la méthode REB, on considère qu'il y a un “ **manque de preuve**” quant à l'efficacité de l'amitriptyline pour le critère de jugement « amélioration globale évaluée par le médecin ».

❖ Selon GRADE

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

Risque de biais : Les quatre études incluses dans notre méta-analyse présentent un risque élevé de biais selon l'outil RoB 2. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, deux études Carette 1995 et Ginsberg 1996 présentent un risque jugé élevé pour 3 domaines du RoB 2 et les deux

autres Don Goldenberg 1996 et de Hanonnen 1998 ont un risque de biais élevé pour deux domaines. De ce fait nous décidons de rétrograder le niveau de preuve de deux niveaux.

Incohérence : Les 4 études suggèrent toutes une différence en faveur de l'amitriptyline par rapport au placebo. Il en résulte une hétérogénéité entre « faible » et « modérée » pour la méta-analyse avec un test I^2 à 38. Cependant on note un p à 0.18 donc non significatif. La valeur « faible » de l'hétérogénéité pourrait donc être due au hasard, et ce d'autant plus que les études incluses dans l'analyse quantitative sont peu nombreuses et de petites tailles. Nous décidons de ce fait de retirer 1 point.

Caractère indirect des résultats : Les critères de jugement des quatre études et leurs populations sont en lien avec l'objectif de notre revue de la littérature. Parmi les 4 études, 2 études mesurent l'efficacité d'une autre molécule en même temps que l'amitriptyline par rapport au placebo. Parmi les 4 études, 2 études ne comparent pas les résultats finaux des patients traités par amitriptyline et par placebo mais observent l'amélioration intragroupe. Nous décidons de ce fait de retirer 1 point.

Imprécision : L'effectif total de notre méta-analyse est de 188 patients et est donc considéré comme faible. Notre différence de moyennes standardisées est à -0.63 avec un intervalle de confiance compris entre [-0.93 ; -0.34]. Elle est donc à moins de 0.5 des limites de l'intervalle de confiance et ses bornes sont cohérentes pour situer l'effet de l'amitriptyline comme étant favorable par rapport au placebo à 95%. Nous décidons donc d'enlever un point au niveau de preuve à cause de l'effectif

Biais de publication : Les études sont toutes à petits échantillons. Le *funnel plot* ne rassemble pas suffisamment d'études pour permettre une interprétation fiable, bien qu'il semble plutôt montrer une distribution symétrique des études. De plus l'analyse quantitative n'inclut que 4 études parmi les 8 identifiées dans l'analyse qualitative. De ce fait, nous ne pouvons exclure un biais de publication. Nous enlevons 1 point au niveau de preuve.

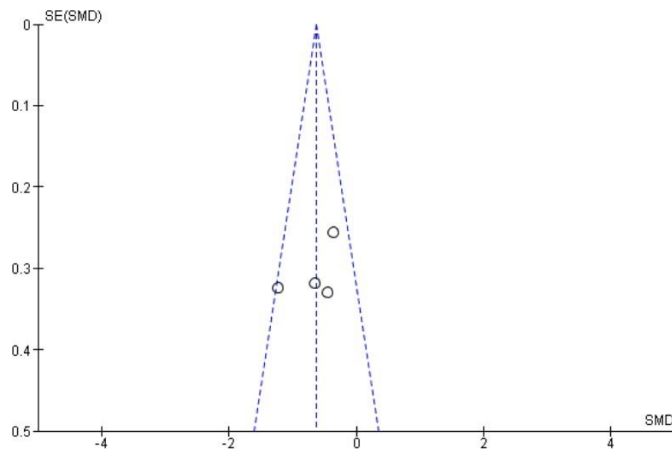


Figure 13 : Représentation du *funnel plot* de l'analyse quantitative pour le critère de jugement « amélioration globale évaluée par le médecin »

Nous n'avons pas rajouté de points pour les items « ampleur de l'effet », « effet dose » et « biais plausibles ».

Conclusion : En prenant en compte ces résultats, nous concluons que l'efficacité de l'amitriptyline est de « très faible » niveau de preuve pour l'amélioration globale de la fibromyalgie évaluée par le médecin.

c) Evaluation de la douleur

❖ Selon REB

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2. Aucune étude n'est concluante pour ce critère de jugement. Moins de la moitié des essais cliniques analysés montrent un résultat statistiquement significatif. Notre méta-analyse de sensibilité, retrouve une différence de moyennes standardisées significative à -0.42 en faveur de l'amitriptyline mais avec une hétérogénéité importante et un test I^2 à 71 %. Cependant notre analyse qualitative inclut seulement 9 études parmi les 11 identifiées lors de l'analyse qualitative. De plus le faible nombre d'études incluses rend l'interprétation du *funnel plot* non fiable, bien que ce dernier semble être plutôt symétrique. Ces éléments ne nous permettent pas d'affirmer qu'il n'existe pas de biais de publication.

Selon la méthode REB, on considère qu'il y a un « manque de preuve » quant à l'efficacité de l'amitriptyline pour le critère de jugement « intensité de la douleur ».

❖ Selon GRADE

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

Risque de biais : Les neuf études incluses dans notre méta-analyse présentent un risque élevé de biais selon l'outil RoB 2. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, l'étude de Kempeneers 1994 a des risques jugés élevés pour les 5 domaines du RoB 2. Les études Carette 1986, 1994, 1995, Ginsberg 1996 et Braz 2012 présentent un risque jugé élevé pour 3 domaines du RoB 2. Les études de Don Goldenberg 1996 et de Hanonnen 1998 ont un risque de biais élevé pour deux domaines et l'étude de Fors 2002 a un seul domaine à risque élevé de biais. De ce fait nous décidons de rétrograder le niveau de preuve de deux niveaux.

Incohérence : Parmi les 9 études éligibles à la méta-analyse, 7 suggèrent une différence en faveur de l'amitriptyline. L'étude Kempeneers 1994 suggère une absence de différence entre l'amitriptyline et le placebo. L'étude Braz 2012 suggère une différence en faveur du placebo. Ceci explique l'hétérogénéité de la méta-analyse qui apparaît comme substantielle avec un test I^2 à 71% avec une valeur de $p = 0.0006$. Nous décidons de ce fait qu'il existe un risque sérieux et nous décidons de rétrograder de deux niveaux.

Caractère indirect des résultats : Les critères de jugement des 9 études et leurs populations sont en lien avec l'objectif de notre revue de la littérature. Parmi les 9 études, 5 études mesurent l'efficacité d'une autre molécule en même temps que l'amitriptyline par rapport au placebo. Parmi les 9 études, 3 études ne comparent pas les résultats finaux des patients traités par amitriptyline et par placebo mais observent l'amélioration intragroupe. Nous décidons de ce fait de retirer 1 point.

Imprécision : L'effectif total de notre méta-analyse est de 452 patients et est donc considéré comme de bonne taille. Notre différence de moyenne standardisée est à -0.42 avec un intervalle de confiance compris entre [-0.79 ; - 0.04]. Elle est donc à moins de 0.5 des limites de l'intervalle de confiance et ses bornes sont cohérentes pour situer l'effet de l'amitriptyline comme étant favorable par rapport au placebo à 95%. Nous décidons donc de ne pas enlever de point.

Biais de publication : Les études sont toutes à petits échantillons. Le *funnel plot* ne rassemble pas suffisamment d'études pour permettre une interprétation fiable, bien qu'il semble plutôt

montrer une distribution symétrique des études. De ce fait, nous ne pouvons exclure un biais de publication. Nous enlevons 1 point au niveau de preuve.

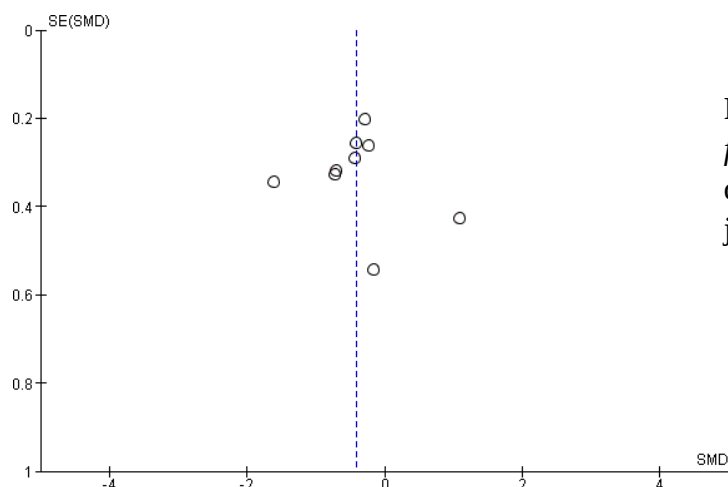


Figure 14 : Représentation du *funnel plot* de l'analyse quantitative pour le critère de jugement « douleur »

Nous n'avons pas rajouté de points pour les items " ampleur de l'effet", " effet dose" et " biais plausibles ".

Conclusion : En prenant en compte ces résultats, nous concluons que l'efficacité de l'amitriptyline est de « très faible niveau de preuve » pour l'intensité de la douleur.

d) Evaluation du Sommeil

❖ Selon REB

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2. Aucune étude n'est concluante pour ce critère de jugement. Moins de la moitié des essais cliniques analysés montrent un résultat statistiquement significatif. Notre méta-analyse de sensibilité, retrouve une différence de moyennes standardisées significative à -0.53 en faveur de l'amitriptyline. L'hétérogénéité est considérée comme modérée avec un test I^2 à 31 %. Cependant notre analyse qualitative inclut seulement 5 études parmi les 9 identifiées lors de l'analyse qualitative. De plus, le faible nombre d'études incluses rend l'interprétation du *funnel plot* non fiable, bien que ce dernier semble être plutôt symétrique. Ces éléments ne nous permettent pas d'affirmer qu'il n'existe pas de biais de publication.

Selon la méthode REB, on considère qu'il y a un “[manque de preuve](#)” quant à l'efficacité de l'amitriptyline pour le critère de jugement « troubles du sommeil ».

❖ Selon GRADE

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

Risque de biais : Les 5 études incluses dans notre méta-analyse présentent un risque élevé de biais selon l'outil RoB 2. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, les études Carette 1995, Ginsberg 1996 et Braz 2012 présentent un risque jugé élevé pour 3 domaines du RoB 2. Les études de Don Goldenberg 1996 et de Hanonnen 1998 ont un risque de biais élevé pour deux domaines. De ce fait nous décidons de rétrograder le niveau de preuve de deux niveaux.

Incohérence : 4 études suggèrent une différence en faveur de l'amitriptyline. L'étude Braz 2012 suggère une différence en faveur du placebo. Le score d'hétérogénéité de la méta-analyse apparaît comme faible avec un test I^2 à 31%, mais un $p = 0.22$ donc non significatif. Cela suggère que la valeur « faible » de l'hétérogénéité pourrait être due hasard (faible nombre d'études incluses, effectifs de petites tailles). Nous décidons de ce fait qu'il existe malgré tout un risque et nous décidons de rétrograder d'un seul niveau.

Caractère indirect des résultats : Les critères de jugement des 5 études et leurs populations sont en lien avec l'objectif de notre revue de la littérature. Parmi les 5 études, 3 études mesurent l'efficacité d'une autre molécule en même temps que l'amitriptyline par rapport au placebo. Parmi les 5 études, 2 études ne comparent pas les résultats finaux des patients traités par amitriptyline et par placebo mais les auteurs observent une amélioration intragroupe. Nous décidons de ce fait de retirer 1 point.

Imprécision : L'effectif total de notre méta-analyse est de 216 patients et est donc considéré comme de faible taille. Notre différence de moyennes standardisées est à -0.53 avec un intervalle de confiance compris entre [-0.80 ; -0.25]. Elle est donc à moins de 0.5 des limites de l'intervalle de confiance et ses bornes sont cohérentes pour situer l'effet de l'amitriptyline comme étant favorable par rapport au placebo à 95%. Nous décidons donc de retirer 1 point au niveau de preuve, à cause de l'effectif

Biais de publication : Les études sont toutes à petits échantillons. Le *funnel plot* ne rassemble pas suffisamment d'études pour permettre une interprétation fiable, bien qu'il semble plutôt montrer une distribution symétrique des études. De plus l'analyse quantitative n'inclut que 5 études parmi les 9 identifiées dans l'analyse qualitative. De ce fait, nous ne pouvons exclure un biais de publication. Nous décidons de retirer 1 point au niveau de preuve.

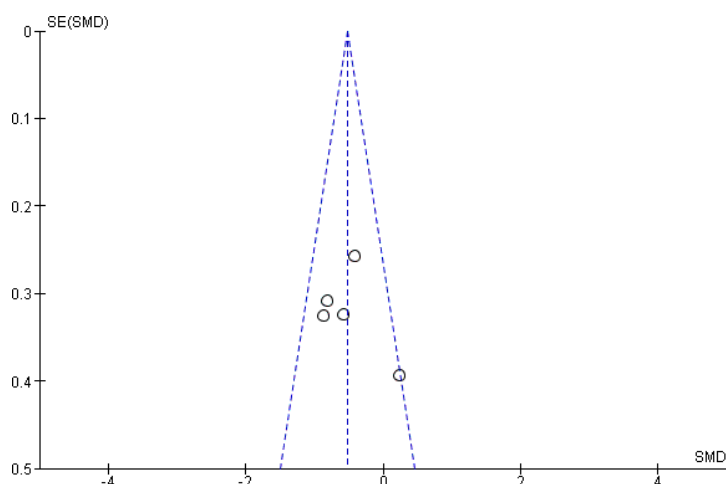


Figure 15 : Représentation du *funnel plot* de l'analyse quantitative pour le critère de jugement « trouble du sommeil »

Nous n'avons pas rajouté de points pour les items « ampleur de l'effet », « effet dose » et « biais plausibles ».

Conclusion : En prenant en compte ces résultats, nous concluons que l'efficacité de l'amitriptyline est de « très faible » niveau de preuve pour le sommeil.

e) Evaluation de la fatigue

❖ Selon REB

Toutes les études incluses sont à risque élevé de biais selon l'outil RoB 2. Aucune étude n'est concluante pour ce critère de jugement. Moins de la moitié des essais cliniques analysés montrent un résultat statistiquement significatif. Notre méta-analyse de sensibilité, retrouve une différence de moyennes standardisées à -0.36 en faveur de l'amitriptyline mais n'est pas significative. L'hétérogénéité est considérée comme importante avec un test I^2 à 65 %. Cependant notre analyse qualitative inclut seulement 6 études parmi les 9 identifiées lors de l'analyse qualitative. De plus le faible nombre d'études incluses rend l'interprétation du *funnel*

plot non fiable, bien que ce dernier semble être plutôt symétrique. Ces éléments ne nous permettent pas d'affirmer qu'il n'existe pas de biais de publication.

Selon la méthode REB, on considère qu'il y a un “manque de preuve” quant à l'efficacité de l'amitriptyline pour le critère de jugement « fatigue ».

❖ Selon GRADE

Le niveau de preuve initial est élevé dans notre cas car toutes nos études sont des essais cliniques randomisés.

Risque de biais : Les 6 études incluses dans notre méta-analyse présentent un risque élevé de biais selon l'outil RoB 2. Nous considérons qu'il existe un risque crucial de biais pour plusieurs critères susceptible d'altérer gravement les résultats. En effet, l'étude de Kempenaers 1994 a des risques jugés élevés pour les 5 domaines du RoB 2. Les études Carette 1995, Ginsberg 1996 et Braz 2012 présentent un risque jugé élevé pour 3 domaines du RoB 2. Les études de Don Goldenberg 1996 et de Hanonnen 1998 ont un risque de biais élevé pour deux domaines. De ce fait nous décidons de rétrograder le niveau de preuve de deux niveaux.

Incohérence : 4 études suggèrent un résultat en faveur de l'amitriptyline. Les études Hanonnen 1998 et Braz 2012 suggère un résultat en faveur du placebo. L'hétérogénéité de la méta-analyse apparaît comme substantielle avec un test I^2 à 65% et une valeur de p à 0.01. Nous décidons de ce fait qu'il existe un risque important et nous décidons de rétrograder de deux niveaux.

Caractère indirect des résultats : Les critères de jugement des 6 études et leurs populations sont en lien avec l'objectif de notre revue de la littérature. Parmi les 6 études, 4 études mesurent l'efficacité d'une autre molécule en même temps que l'amitriptyline par rapport au placebo. Parmi les 6 études, 3 études ne comparent pas les résultats finaux des patients traités par amitriptyline et par placebo mais observent l'amélioration intragroupe.

Nous décidons de ce fait de retirer 1 point.

Imprécision : L'effectif total de notre méta-analyse est de 230 patients. Il est donc considéré comme de faible taille. Notre différence de moyennes standardisées est non significative ($p = 0.13$) de -0.36 avec un intervalle de confiance à 95% compris entre [-0.82 ; 0.11]. Bien que l'intervalle de confiance soit restreint, il est incohérent car il comprend des valeurs en faveur et en défaveur de l'amitriptyline. Nous décidons donc de retirer 2 points au niveau de preuve.

Biais de publication : Les études sont toutes à petits échantillons. Le *funnel plot* ne rassemble pas suffisamment d'études pour permettre une interprétation fiable, bien qu'il semble plutôt montrer une distribution symétrique des études. De ce fait, nous ne pouvons exclure un biais de publication. Nous rétrogradons d'un niveau le niveau de preuve.

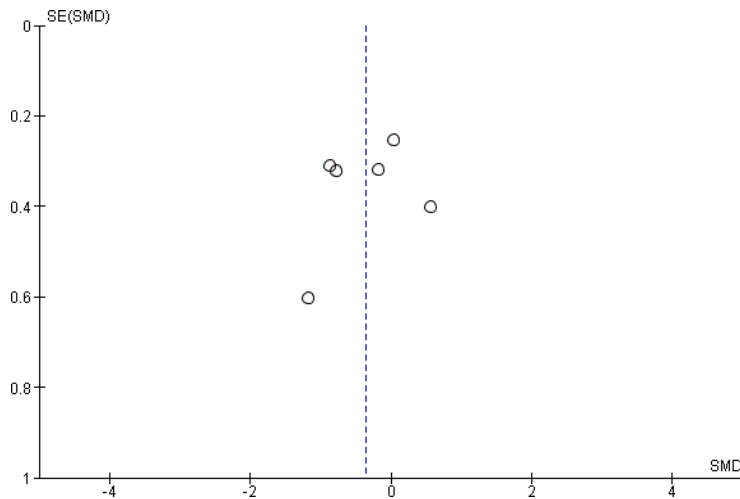


Figure 16 : Représentation du *funnel plot* de l'analyse quantitative pour le critère de jugement « fatigue »

Nous n'avons pas rajouté de points pour les items " ampleur de l'effet", "effet dose" et " biais plausibles ".

Conclusion : En prenant en compte ces résultats, nous concluons que l'efficacité de l'amitriptyline est de « très faible » niveau de preuve pour la fatigue

V- DISCUSSION

1. Interprétation des résultats :

A l'issue d'une recherche systématique, 12 essais cliniques randomisés comparant l'amitriptyline à un placebo dans le cadre du traitement des symptômes de la fibromyalgie ont été inclus et analysés. L'étude (Jaeschke 1991) n'a pu être incluse ni dans les analyses qualitatives ni dans les analyses quantitatives pour des raisons de méthodologie et de résultats manquants, ce qui a conduit à son exclusion secondaire. La fibromyalgie étant un syndrome, les études incluses mesurent toutes au moins 1 des critères de jugement d'intérêt suivant : amélioration globale évaluée par le patient, amélioration globale évaluée par le médecin, douleur, trouble du sommeil, fatigue. Cet ensemble de 12 études regroupe approximativement 334 patients traités par amitriptyline avec des doses allant de 25 mg à 50 mg, et 293 patients traités par placebo, soit environ 627 patients (nombre de patients estimé à partir des données des articles analysés en per protocole ou en intention de traiter modifiée). Notre population est constituée principalement de femmes, provenant d'Europe, d'Amérique du Nord et du Sud, avec un suivi allant de 4 à 32 semaines et une durée de traitement par amitriptyline testée entre 4 semaines et 6 mois. Etant donné que toutes les études incluses sont à haut niveau de risques de biais, seule une analyse quantitative de sensibilité a pu être réalisée pour chaque critère de jugement.

L'analyse quantitative de chaque critère de jugement, n'a été réalisée que sur une partie des études incluses dans l'analyse qualitative. En effet, les résultats de certaines études n'étaient pas exploitables, malgré le fait que l'on ait contacté leurs auteurs et reçu de l'aide de Mme Veloso. De plus, le nombre de patients n'a pas pu être déterminé avec précision car certaines études ne précisent pas les effectifs par groupe.

Nous constatons par ailleurs qu'une seule étude sur les 12 analysées, trouve systématiquement un résultat non significatif en faveur du placebo pour chacun des critères de jugement mesurés (Braz 2012). Bien que ce soit l'essai clinique le plus récent de notre méta-analyse, sa

méthodologie est jugée à haut risque de biais (comme toutes les autres études incluses). De plus, les effets indésirables n'ont pas été rapportés par les auteurs.

Nous observons quelques écarts en termes de significativité des résultats des études entre les analyses qualitatives et quantitatives.

Ainsi, pour le critère de jugement - amélioration globale par le patient, ressenti du patient- on constate que la comparaison des résultats finaux entre l'amitriptyline et le placebo est non significative pour l'étude Carette 1995, alors que les auteurs de l'article rapportent la différence comme étant significative. Nous expliquons cela par une possible imprécision dans le calcul des bornes de l'intervalle de confiance (borne inférieure de l'intervalle de confiance à 0.01 dans notre analyse quantitative).

Pour le critère de jugement – amélioration globale par le patient, répercussion de la pathologie sur sa qualité de vie - nous retrouvons dans notre analyse quantitative un résultat significatif : Heymann 2001, alors que les auteurs expriment dans l'article que l'amélioration n'est pas significative par rapport au placebo ($p_{\text{time}} < 0.001$ et $p_{\text{group} \times \text{time}} = 0.253$). Nous expliquons la différence de significativité par le fait que nous avons comparé les résultats finaux et non les améliorations.

Pour le critère de jugement - douleur – on constate dans l'analyse quantitative un résultat significatif en faveur du placebo pour l'étude Braz 2012, alors que les auteurs rapportent dans leur article que les améliorations entre les 3 groupes de traitements sont non significatives. Nous supposons que cette différence d'interprétation des résultats vient de ce que nous comparons les résultats finaux des groupes de traitements et non leurs améliorations.

Pour le critère de jugement – sommeil – on constate dans l'analyse quantitative un résultat non significatif pour Hanonnen 1998, alors que les auteurs notent que l'index de sommeil est statistiquement significatif par rapport aux autres groupes de traitements ($p < 0.01$), et que le tableau de résultat de l'article ne rapporte que la significativité des résultats en intra groupe. Nous supposons que la divergence dans l'interprétation des résultats vient de ce que nous comparons les résultats finaux entre amitriptyline et placebo, ce qui ne semble pas être le cas des auteurs.

Pour le critère de jugement – fatigue – les résultats sont globalement concordants. L'analyse quantitative confirme la très faible différence d'effet pour Don Goldenberg 1996 en faveur de l'amitriptyline et pour Hanonnen 1998 en faveur du placebo.

A noter que pour les articles qui n'avaient pas directement comparé les résultats finaux des traitements par amitriptyline et par placebo, la méta-analyse montre :

- Pour Kempenaers 1994 : pas de différence significative ni pour la douleur, ni pour la fatigue
- Pour Don Goldenberg 1996 : un résultat significatif uniquement pour la douleur, en faveur de l'amitriptyline
- Pour Ginsberg 1996 : des résultats significatifs pour l'amélioration globale évaluée par le médecin et le ressenti du patient, la douleur, le sommeil et la fatigue

a) Critère de jugement « amélioration globale évaluée par le patient »

Pour le critère de jugement « amélioration globale évaluée par le patient », il nous a paru pertinent de l'analyser de 2 façons complémentaires : le ressenti global par le patient d'une part et les incapacités dues à la pathologie d'autre part.

L'analyse qualitative du « ressenti global par le patient » regroupe 10 études soit approximativement 552 patients, avec environ 299 patients dans le groupe amitriptyline et 253 dans celui traité par placebo : Carette 1986, Don Goldenberg 1986, Scudds 1989, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Heymann 2001. Les études suggèrent toutes une amélioration des patients avec l'amitriptyline par rapport au placebo, dont 4 avec des résultats statistiquement significatifs. Selon la méthodologie REB, aucune étude n'est concluante, et moins de 50% des résultats en faveur de l'amitriptyline sont significatifs. Donc l'analyse qualitative conclut à un manque de preuve de l'efficacité de l'amitriptyline pour ce critère de jugement.

La méta-analyse n'a pu être réalisée qu'à partir de 4 de ces études : Carette 1995, Don Goldenberg 1996, Ginsberg 1996 et Hanonnen 1998. Elle regroupe 99 patients dans le groupe amitriptyline et 91 patients dans le groupe placebo, soit 190 patients. Son résultat montre une amélioration significative en faveur de l'amitriptyline ($p = 0.003$) avec une taille d'effet modérée (SMD = -0.68) et un intervalle de confiance à 95% concordant compris entre [-1.13,

-0.24]. L'hétérogénéité est de $I^2 = 55\%$, $p = 0.08$ ce qui est jugé soit « élevé » selon la méthode REB, soit « modéré » selon le GRADE. Le niveau de preuve global de l'efficacité de l'amitriptyline par rapport au placebo pour ce critère de jugement est estimé comme étant « un manque de preuve » selon la méthode REB et de « très faible niveau de preuve » selon le GRADE.

L'analyse qualitative des « répercussions globales de la pathologie sur la qualité de vie du patient » regroupe 3 études soit 136 patients, avec 71 patients dans le groupe amitriptyline et 65 dans celui traité par placebo : Don Goldenberg 1996, Heymann 2001, Braz 2012. Les analyses qualitatives et quantitatives regroupent les mêmes études pour les mêmes critères de jugement. Elles comprennent 2 études suggérant une amélioration des patients traités par amitriptyline par rapport à ceux traités par placebo, mais aucune n'a de résultat significatif, et une étude ne montrant pas de différence entre les 2 molécules. Les résultats des 3 essais sont exploratoires. Selon la méthode REB le niveau de preuve de l'analyse qualitative est « manque de preuve »

La méta-analyse inclut les mêmes études : Don Goldenberg 1996, Heymann 2001, Braz 2012. Elle regroupe 71 patients dans le groupe amitriptyline et 65 patients dans le groupe placebo, soit 136 patients. Le résultat montre une amélioration non significative en faveur de l'amitriptyline ($p = 0.21$) avec un intervalle de confiance à 95% compris entre $[-15.22 ; 3.27]$ incohérent et imprécis et une hétérogénéité jugée « considérable » par le GRADE et « important » par le REB : $I^2 = 90\%$ avec $p < 0.0001$. Le niveau de preuve global de l'efficacité de l'amitriptyline par rapport au placebo est jugé comme étant « un manque de preuve » selon la méthode REB et de « très faible niveau de preuve » selon le GRADE.

Ainsi, pour le critère de jugement « amélioration globale évaluée par le patient », les résultats des analyses qualitatives et quantitatives sont concordants et suggèrent une différence en faveur du traitement par l'amitriptyline par rapport au placebo. Elle est significative sur le ressenti avec un effet modéré et non significative sur la mesure des répercussions sur la qualité de vie, avec une hétérogénéité de résultats importante entre les études. L'évaluation du niveau de preuve de la méta-analyse conclut à un « manque de preuve » sur l'efficacité de l'amitriptyline par rapport au placebo par la méthode REB et à un « très faible niveau de preuve » par la méthodologie GRADE pour le critère de jugement « amélioration globale évaluée par le patient ».

b) Critère de jugement « amélioration globale évaluée par le médecin »

L'analyse qualitative du critère de jugement « amélioration globale évaluée par le médecin » regroupe 8 études soit approximativement 405 patients, avec environ 224 patients dans le groupe amitriptyline et 181 dans celui traité par placebo : Carette 1986, Don Goldenberg 1986, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996 et Hanonnen 1998. Les 8 études suggèrent une amélioration des patients traités par amitriptyline par rapport à ceux traités par placebo, dont 3 avec un résultat statistiquement significatif et parmi elles, un résultat confirmatoire (Hanonnen 1998). Selon la méthodologie REB, seule l'étude Hanonnen 1998 apporte un résultat confirmatoire mais non concluant (haut niveau de risque de biais de l'étude), et moins de 50% des résultats sont significatifs en faveur de l'amitriptyline. Donc l'analyse qualitative conclut à un manque de preuve de l'efficacité de l'amitriptyline pour ce critère de jugement.

La méta-analyse n'a pu être réalisée qu'à partir de 4 de ces études : Carette 1995, Don Goldenberg 1996, Ginsberg 1996 et Hanonnen 1998. Elle regroupe 98 patients dans le groupe amitriptyline et 90 patients dans le groupe placebo, soit 188 patients. Son résultat montre une amélioration significative en faveur de l'amitriptyline ($p < 0.0001$) avec une taille d'effet modérée ($SMD = -0.63$) et un intervalle de confiance à 95% compris entre $[-0.93 ; -0.34]$ cohérent. L'hétérogénéité est de $I^2 = 38\%$ ce qui est considéré soit comme « modéré » selon la méthode REB, soit comme étant entre « faible » et « modéré » selon le GRADE. Cependant le test d'hétérogénéité est non significatif avec un p relativement important : $p = 0.18$, ce qui laisse un doute sur l'homogénéité réelle du résultat (et non liée au hasard), et ce d'autant plus que les études sont peu nombreuses et de petites tailles.

Ainsi, les résultats de l'étude qualitative suggèrent une différence en faveur de l'amitriptyline avec une homogénéité dans les résultats des études la composant. Elle est concordante avec l'analyse quantitative qui montre une différence significative en faveur d'un traitement par l'amitriptyline par rapport au placebo avec un effet et une hétérogénéité tout deux modérés. L'évaluation du niveau de preuve de la méta-analyse conclut à un « manque de preuve » sur l'efficacité de l'amitriptyline par rapport au placebo par la méthode REB et à un « très faible niveau de preuve » par la méthodologie GRADE pour le critère de jugement « amélioration globale évaluée par le médecin ».

c) Critère de jugement « intensité de la douleur »

L'analyse qualitative du critère de jugement « intensité de la douleur » regroupe 11 études soit approximativement 557 patients, avec environ 297 patients dans le groupe amitriptyline et 260 dans celui traité par placebo : Carette 1986, Don Goldenberg 1986, Scudds 1989, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Fors 2002 et Braz 2012. 9 études suggèrent une amélioration des patients traités par amitriptyline par rapport à ceux traités par placebo, dont 3 avec un résultat statistiquement significatif. 1 étude suggère une absence de différence entre le traitement par amitriptyline ou par placebo (Kempenaers 1994) et 1 étude rapporte un résultat non significatif, en faveur du placebo (Braz 2012). Aucune étude ne présente de résultat confirmatoire pour ce critère de jugement. Selon la méthodologie REB, aucune étude n'est confirmatoire, et moins de 50% des résultats sont significatifs en faveur de l'amitriptyline. Donc l'analyse qualitative conclut à un manque de preuve de l'efficacité de l'amitriptyline pour ce critère de jugement.

La méta-analyse n'a pu être réalisée qu'à partir de 9 de ces études : Carette 1986, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Fors 2002 et Braz 2012. Elle regroupe 245 patients dans le groupe amitriptyline et 207 patients dans le groupe placebo, soit 452 patients. Son résultat montre une amélioration significative en faveur de l'amitriptyline ($p = 0.03$) avec une taille d'effet faible ($SMD = -0.42$) et un intervalle de confiance à 95% compris entre $[-0.79 ; -0.04]$ cohérent. L'hétérogénéité est de $I^2 = 71\%$, $p = 0.0006$ ce qui est jugé soit « important » selon la méthode REB, soit « substantiel » selon le GRADE.

Ainsi, les résultats de l'étude qualitative suggèrent une différence en faveur de l'amitriptyline avec une hétérogénéité dans les résultats des études la composant. Elle est concordante avec l'analyse quantitative qui montre une différence significative en faveur d'un traitement par l'amitriptyline par rapport au placebo avec un effet faible et une hétérogénéité importante. L'évaluation du niveau de preuve de la méta-analyse conclut à un « manque de preuve » sur l'efficacité de l'amitriptyline par rapport au placebo par la méthode REB et à un « très faible niveau de preuve » par la méthodologie GRADE pour le critère de jugement « douleur ».

d) Critère de jugement « troubles du sommeil »

L'analyse qualitative du critère de jugement « troubles du sommeil » regroupe 9 études soit approximativement 433 patients, avec environ 238 patients dans le groupe amitriptyline et 195 dans celui traité par placebo : Carette 1986, Don Goldenberg 1986, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998 et Braz 2012. 7 études suggèrent une amélioration des patients traités par amitriptyline par rapport à ceux traités par placebo, dont 5 avec un résultat statistiquement significatif. 1 étude suggère une absence de différence entre le traitement par amitriptyline ou par placebo (Kempenaers 1994) et 1 étude rapporte un résultat non significatif, en faveur du placebo : Braz 2012. Aucune étude ne présente de résultat confirmatoire pour ce critère de jugement. Selon la méthodologie REB, aucune étude n'est confirmatoire, et moins de 50% des résultats sont significatifs en faveur de l'amitriptyline. Donc l'analyse qualitative conclut à un manque de preuve de l'efficacité de l'amitriptyline pour ce critère de jugement.

La méta-analyse n'a pu être réalisée qu'à partir de 5 études : Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998 et Braz 2012. Elle regroupe 112 patients dans le groupe amitriptyline et 104 patients dans le groupe placebo, soit 216 patients. Son résultat montre une amélioration significative en faveur de l'amitriptyline ($p = 0.0002$) avec une taille d'effet modérée ($SMD = -0.53$) et un intervalle de confiance à 95% compris entre $[-0.80 ; -0.25]$ cohérent. L'hétérogénéité est de $I^2 = 31\%$ ce qui est considéré comme « faible » selon la méthode REB et la méthode GRADE. Cependant le test d'hétérogénéité est non significatif avec un p relativement important : $p = 0.22$, ce qui laisse un doute sur l'homogénéité réelle du résultat et ce d'autant plus que les études sont peu nombreuses et de petites tailles.

L'étude qualitative et l'étude quantitative diffèrent quant à l'appréciation de l'hétérogénéité (études aux résultats contradictoires dans l'analyse qualitative versus bonne homogénéité calculée dans l'analyse quantitative, cependant pondérée par un p élevé). Cependant les résultats de l'étude qualitative suggèrent une différence en faveur de l'amitriptyline et est concordante avec l'analyse quantitative qui montre une différence significative en faveur d'un traitement par l'amitriptyline par rapport au placebo avec un effet modéré. L'évaluation du niveau de preuve de la méta-analyse conclut à un « manque de preuve » sur l'efficacité de l'amitriptyline par rapport au placebo par la méthode REB et à un « très faible niveau de preuve » par la méthodologie GRADE pour le critère de jugement « sommeil ».

e) Critère de jugement « fatigue »

L'analyse qualitative du critère de jugement « fatigue » regroupe 8 études soit approximativement 374 patients, avec environ 211 patients dans le groupe amitriptyline et 163 dans celui traité par placebo : Don Goldenberg 1986, Carette 1994, Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998 et Braz 2012. 5 études suggèrent une amélioration des patients traités par amitriptyline par rapport à ceux traités par placebo, dont 3 avec un résultat statistiquement significatif. 1 étude suggère une absence de différence entre l'amitriptyline et le placebo (Don Goldenberg 1996). 2 études rapportent un résultat non significatif, en faveur du placebo : Hanonnen 1998 et Braz 2012. Aucune étude ne présente de résultat confirmatoire pour ce critère de jugement. Selon la méthodologie REB, aucune étude n'est confirmatoire, et moins de 50% des résultats sont significatifs en faveur de l'amitriptyline. Donc l'analyse qualitative conclut à un manque de preuve de l'efficacité de l'amitriptyline pour ce critère de jugement.

La méta-analyse n'a pu être réalisée qu'à partir de 6 de ces études : Kempenaers 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998 et Braz 2012. Elle regroupe 118 patients dans le groupe amitriptyline et 112 patients dans le groupe placebo, soit 230 patients. Son résultat montre une amélioration non significative en faveur de l'amitriptyline ($p = 0.13$) avec une taille d'effet faible ($SMD = -0.36$) et un intervalle de confiance à 95% compris entre $[-0.82 ; 0.11]$ incohérent. En effet il contient des valeurs de résultat en faveur et en défaveur de l'amitriptyline. L'hétérogénéité est de $I^2 = 65\%$ avec $p = 0.01$ ce qui est jugé soit « important » selon la méthode REB, soit « substantiel » selon le GRADE.

Ainsi, pour le critère de jugement « fatigue », les analyses qualitatives et quantitatives sont concordantes dans leurs résultats et leur hétérogénéité. Elles suggèrent une efficacité de l'amitriptyline par rapport au placebo, mais sans être significatives, avec une hétérogénéité importante, et pour un effet, si existant, estimé faible. Le niveau de preuve global de l'efficacité de l'amitriptyline par rapport au placebo est jugé comme étant « un manque de preuve » selon la méthode REB et de « très faible niveau de preuve » selon le GRADE pour ce critère de jugement.

f) Pour les effets indésirables et les perdus de vue

– Les effets indésirables :

L'analyse qualitative des effets indésirables inclut 6 études soit approximativement 444 patients, avec environ 241 patients dans le groupe amitriptyline et 203 dans celui traité par placebo : Carette 1986, Don Goldenberg 1986, Carette 1994, Ginsberg 1996, Hanonnen 1998 et Heymann 2001.

En ce qui concerne la fréquence des effets indésirables : 3 études suggèrent une augmentation de la fréquence des effets indésirables chez les patients traités par amitriptyline, dont 1 étude avec un résultat significatif. 2 études ne retrouvent pas de différence entre l'amitriptyline et le placebo en termes de fréquence d'effets indésirables. 1 étude retrouve un nombre d'effets indésirables plus important dans le groupe placebo que dans le groupe amitriptyline. Aucune étude ne présente de résultat confirmatoire pour ce critère de jugement.

On constate une grande hétérogénéité dans les résultats.

En ce qui concerne la nature des effets indésirables, aucune analyse n'est réalisable, ni qualitative ni quantitative, du fait d'un manque de données. Les auteurs constatent dans 4 études, un nombre plus important d'effets indésirables spécifiques de l'amitriptyline (sécheresse buccale / somnolence) dans le groupe traité par la molécule active que dans le groupe placebo. Dans 1 étude, les auteurs ne donnent aucune précision sur la nature des effets indésirables. Dans 1 étude, les patients ne rapportent pas plus d'effets indésirables spécifiques de l'amitriptyline que dans le groupe placebo.

La méta-analyse n'a pu être réalisée qu'à partir de 3 de ces études : Carette 1986, Ginsberg 1996, et Heymann 2001. Elle regroupe 91 patients dans le groupe amitriptyline et 94 patients dans le groupe placebo, soit 185 patients. Son résultat montre une tendance non significative à un nombre plus important d'effets indésirables ($p = 0.31$) avec une taille d'effet immense ($SMD = 2.93$) mais associée à un intervalle de confiance à 95% très étendu donc très imprécis compris entre $[0.37 ; 23.42]$ quoique cohérent. L'hétérogénéité est de $I^2 = 92\%$ avec $p < 0.00001$ ce qui est jugé soit « important » selon la méthode REB, soit « considérable » selon le GRADE.

– Les perdus de vue :

Etant donné que toutes les études ayant pu être analysées pour l'analyse qualitative ont été incluses dans l'analyse quantitative, il ne nous a pas semblé pertinent de réaliser une analyse qualitative.

La méta-analyse a pu être réalisée à partir de 10 études : Carette 1986, Scudds 1989, Carette 1994, Carette 1995, Don Goldenberg 1996, Ginsberg 1996, Hanonnen 1998, Heymann 2001, Fors 2002 et Braz 2012. Elle regroupe 351 patients dans le groupe amitriptyline et 312 patients dans le groupe placebo, soit 663 patients. Son résultat suggère une augmentation non significative du nombre de perdus de vue dans le groupe placebo par rapport au groupe amitriptyline ($p = 0.05$, intervalle de confiance à 95% compris entre [0.48 ; 1]). L'hétérogénéité est de $I^2 = 0\%$ avec $p = 0.72$, ce qui est jugé « faible » par les méthodes REB et GRADE. Cependant le test d'hétérogénéité est non significatif avec un p relativement important : $p = 0.72$, ce qui laisse un doute sur l'homogénéité réelle du résultat, et ce d'autant plus que les études incluses sont de petites tailles.

Ainsi, pour la fréquence des effets indésirables, l'analyse qualitative et l'analyse quantitative présentent toutes deux une grande hétérogénéité. De plus, l'intervalle de confiance est très large et comprend des valeurs en faveur et en défaveur de l'amitriptyline. Ces résultats suggèrent une fréquence d'effets indésirables plus importante dans le groupe traité par amitriptyline, sous réserve du faible nombre d'études et de patients inclus dans la méta-analyse. La lecture des articles suggère quant à elle que les effets spécifiques de l'amitriptyline (sécheresse buccale et somnolence) sont identifiables par rapport au placebo.

Cependant, l'adhésion des patients au traitement par amitriptyline a tendance à être meilleure que dans le groupe placebo, mais de façon non significative. Ceci laisse à penser que soit les effets indésirables restent tolérables par rapport aux bénéfices perçus par les patients, soit les patients traités par amitriptyline ont bénéficié d'un suivi différent de la part du personnel soignant (levée du double aveugle et écarts au protocole).

En résumé : Le point le plus important est le fait que pour chacun de nos critères de jugement, le niveau de preuve est systématiquement jugé comme « très faible » et « manque de preuve », ce qui sont les plus bas niveaux respectifs de chacune des méthodes utilisées : GRADE et REB.

Ceci est dû à la mauvaise qualité méthodologique des études incluses, toutes étant estimées à haut niveau de risque de biais. Il convient donc d'interpréter les résultats de nos différents critères de jugement avec prudence, et de considérer que l'efficacité de l'amitriptyline n'est établie pour aucun domaine d'intérêt. Cependant, notre étude est actuellement celle qui a la plus grande taille, que ce soit en termes de nombre d'études ou de patients inclus, et les résultats de nos critères de jugement sont systématiquement en faveur de l'amitriptyline, avec un effet estimé faible à modéré suggérant un effet bénéfique de la molécule dans la prise en charge des symptômes de la pathologie. De plus, bien que les effets indésirables soient rapportés de façon plus fréquente dans le groupe amitriptyline que dans le groupe placebo, l'adhésion des patients au traitement a tendance à être meilleure dans le groupe avec le traitement actif. Ceci pourrait s'expliquer soit par le fait que les effets indésirables restent tolérables par rapport aux bénéfices perçus par les patients, soit par une levée du double aveugle du fait des effets indésirables spécifiques de la molécule et d'une différence dans la prise en charge, non protocolaire, des patients par l'équipe soignante.

2. Confrontation à la littérature :

3 méta-analyses dont l'objet est « l'efficacité de l'amitriptyline dans la fibromyalgie », ont été identifiées :

- *Amitriptyline for fibromyalgia in adults*_de_R Andrew Moore, Sheena Derry, Dominic Aldington, Peter ColePhilip J Wiffen Authors' declarations of interest, Version published: 31 July 2015 (5).
- Häuser W, Petzke F, Üçeyler N, Sommer C. *Comparative efficacy and acceptability of amitriptyline, duloxetine and milnacipran in fibromyalgia syndrome: a systematic review with meta-analysis. Rheumatology.* 2011;50(3):532-43 (29).
- Nishishinya B, Urrútia G, Walitt B, et al. *Amitriptyline in the treatment of fibromyalgia: a systematic review of its efficacy. Rheumatology.* 2008;47:1741-1746 (30)

Notre méta-analyse se distingue des travaux de recherche précédents par :

- L'inclusion uniquement d'essais cliniques randomisés comparant l'amitriptyline à un placebo
- L'absence d'exclusion des études de petite taille (pas de nombre minimal de participants)
- L'utilisation des outils les plus récents pour l'évaluation du risque de biais : le ROB 2 publié en août 2019 (11) et une évaluation par deux méthodes robustes de la preuve scientifique de nos résultats : méthode REB (14) et méthode GRADE (15)
- La prise en compte de façon indépendante de 5 critères cliniquement pertinents : 3 symptômes majeurs de la fibromyalgie : la douleur, les troubles du sommeil, la fatigue, ainsi que l'évaluation globale de l'efficacité du traitement par le patient et par le médecin évaluateur
- Nos bases de recherche : ClinicalTrials.gov, CENTRAL et MEDLINE (nous n'avons pas pu avoir accès à EMBASE)
- Le nombre d'études incluses dans notre analyse

Les études incluses dans les 3 méta-analyses sont concordantes avec notre travail de recherche. On retrouve dans les 3 méta-analyses et dans la nôtre, les essais cliniques ci-après : [Carette 1986](#) ; [Carette 1994](#) ; [Carette 1995](#) ; [Ginsberg 1996](#) ; [Don Goldenberg 1986](#) ; [Don Goldenberg 1996](#) ; [Hanonnen 1998](#). Les méta-analyses de Häuser 2011 et Nishishinya 2008 sont trop anciennes pour avoir inclus Braz 2012. Cette dernière a été incluse dans notre travail de recherche et dans la méta-analyse de Moore 2015. L'étude Jaeschke 1991 a été exclue des 3 méta-analyses à cause de sa méthodologie particulière et elle a finalement aussi été exclue secondairement de notre méta-analyse. La méta-analyse de Moore 2015 évalue principalement la douleur et n'a donc pas inclus les études [Fors 2002](#) ; [Heymann 2001](#) et [Scudds 1989](#). Les auteurs ont aussi pris comme critère d'exclusion les études comportant moins de 10 participants par bras de traitement, ce qui les a conduits à exclure l'étude Kempenaers 1994. Ces 4 essais sont inclus dans les 2 autres méta-analyses et dans la nôtre. A noter que l'étude Moore 2015 inclut un autre article : Zanette 2014 (31) qui compare l'amitriptyline à la mélatonine sans placebo. Cet article ne correspond pas à nos critères d'inclusion.

L'évaluation de la qualité des études est discordante selon les méta-analyses. Nous constatons que les études sont considérées comme étant d'un niveau méthodologique « médiocre » à « modéré » dans Häuser 2011 et « élevé » dans Nishishinya 2008 ce qui est relativement

concordant. A l'opposé, l'évaluation du risque de biais pour chaque étude a été systématiquement estimé comme important dans la méta-analyse de Moore 2015 ainsi que dans la nôtre. Dans les 4 méta-analyses dont la nôtre, les auteurs ont respecté le fait d'analyser le risque de biais de façon indépendante entre 2 auteurs puis la mise en commun et la résolution des divergences par la discussion avec un troisième examinateur. Cependant la méthode utilisée pour l'évaluation des biais diffère : dans la méta-analyse la plus ancienne (Nishishinya 2008) les auteurs ont utilisé l'échelle de Schulz (32) et le score de Jadad (33). Les 2 autres études et notre travail de recherche utilisent l'outil de Cochrane Collaboration (11, 34, 35), mais avec différentes versions : Häuser 2011 utilise la version 5.01 parue entre février 2008 et septembre 2009 (34), alors que Moore 2015 utilise la version 5.1 de mars 2011 (35). Nous avons utilisé la version la plus récente qui propose l'utilisation de l'outil RoB 2 (11). Cette évolution dans les outils d'analyse des biais explique la divergence observée entre les méta-analyses anciennes (Nishishinya 2008 et Häuser 2011) et récentes (Moore 2015 et la nôtre) quant à l'évaluation du risque de biais des études incluses.

Nos résultats sont globalement concordants avec ceux des autres méta-analyses. En effet, en ce qui concerne la douleur, notre méta-analyse et les 3 autres trouvent une amélioration significative des patients traités par amitriptyline par rapport à ceux traités par placebo. L'effet est estimé comme étant faible pour notre méta-analyse ainsi que pour celle de Häuser 2011. Dans la méta-analyse de Moore 2015 les auteurs supposent que leurs résultats sont surestimés, et trouvent que 36% des patients présentent une réduction d'au moins 50% de l'intensité de leur douleur initiale sous traitement par amitriptyline, ce qui correspond à nombre de patients à traiter de 1 sur 4. Les autres critères de jugement de notre étude n'ont pas été mesurés par Moore 2015. Nos résultats pour les critères de jugement « trouble du sommeil » et « amélioration globale évaluée par le médecin » sont concordants avec les résultats des études de Nishishinya 2008 et de Häuser 2011 : résultats significatifs en faveur de l'amitriptyline. Nous avons observé le critère de jugement « amélioration globale par le patient » selon 2 points de vue : « ressenti du patient » et « impact de la pathologie sur la qualité de vie du patient ». Les résultats « ressenti du patient » sont significativement favorables à l'amitriptyline dans notre étude, ce qui est concordant avec l'analyse qualitative de Nishishinya 2008. Les résultats de « l'impact de la pathologie sur la qualité de vie du patient » sont concordants avec ceux de l'analyse qualitative de Häuser 2011 : non significatif. Nos résultats pour le critère de jugement « fatigue » sont moins probants que ceux obtenus dans les 2 méta-analyses Nishishinya 2008

et de Häuser 2011. Nous trouvons une tendance non significative à l'amélioration par l'amitriptyline, avec un effet, si réel, estimé faible. A l'inverse les auteurs des 2 méta-analyses ont trouvé une amélioration significative de l'amitriptyline (sans estimation de l'effet pour l'étude Nishishinya 2008, et avec une estimation d'effet comme étant faible pour Häuser 2011). Cela peut s'expliquer par le fait que les 2 méta-analyses n'incluent pas l'étude de Braz 2012, qui apporte de l'hétérogénéité et une diminution de l'estimation de l'effet pour l'ensemble des critères de jugement ci-après : douleur, fatigue, troubles du sommeil, amélioration évaluée par le médecin examinateur, et FIQ.

L'évaluation du niveau de preuve des critères de jugement : amélioration globale du patient mesurée par ce dernier ou par le médecin examinateur, douleur, troubles du sommeil et fatigue sont jugés dans notre méta-analyse comme étant comme « très faible » selon le GRADE et « manque de preuve » selon la méthode REB. Cette analyse des résultats remet en question l'efficacité de l'amitriptyline pour ces critères de jugement, même lorsque les résultats sont significatifs dans l'analyse quantitative et sont concordants avec l'analyse qualitative. L'étude Moore 2015 conclut de la même manière pour son critère de jugement « douleur », se positionnant ainsi comme la première méta-analyse à remettre en question l'efficacité de l'amitriptyline, et en alertant sur la probable surestimation de son effet réel. En effet, l'amitriptyline a été considérée dans les 2 méta-analyses les plus anciennes comme un gold standard, et par de nombreux médecins de l'époque comme une molécule de première intention. Ce changement dans l'estimation de l'efficacité de l'amitriptyline, s'explique par la prise en compte du faible nombre d'études et de patients inclus dans les méta-analyses, ainsi que par la réévaluation du risque de biais des études incluses grâce à de nouveaux outils.

En résumé : Notre étude est globalement concordante au niveau des études incluses et des résultats avec les autres méta-analyses sur ce sujet. Le changement dans la méthodologie utilisée pour l'évaluation du risque de biais des études incluses conduit cependant à une interprétation des résultats très différente de celle des méta-analyses plus anciennes. Ceci est concordant avec la méta-analyse de Moore 2015 qui, sans remettre en question l'existence même d'une efficacité de l'amitriptyline pour la douleur, alerte sur la surestimation de cet effet, et insiste sur le fait que l'effet est attendu sur un nombre limité de patients (au maximum, un quart). Notre travail permet d'étendre cette remise en question de l'efficacité de l'amitriptyline

au traitement d'autres symptômes : en particulier la fatigue et à l'amélioration globale évaluée par le patient, mais aussi dans une moindre mesure, aux troubles du sommeil, à la douleur et à l'amélioration globale du patient évaluée par le médecin.

3. Les limites de notre étude :

Nous avons rencontré plusieurs limites dans la réalisation de cette revue de la littérature :

- Peu d'études ont été incluses dans notre revue de la littérature. Cela s'explique par le fait que nous comparons l'amitriptyline à un placebo, or de nombreuses études comparent l'amitriptyline uniquement à une molécule active. Le plus souvent dans ces études l'amitriptyline est considérée comme le gold standard.
- La méta-analyse n'a pu porter sur toutes les études incluses traitant du critère de jugement du fait de résultats rapportés sous une forme non exploitable. Nous avons essayé de joindre les différents auteurs pour récupérer les données brutes mais aucun auteur n'a répondu à notre demande favorablement.
- Les auteurs ont utilisé des unités de mesures et des échelles différentes pour leur critère de jugement, mais également des scores qui étaient non reproductibles entre les études. Ces différences rendent l'interprétation des résultats un peu plus approximative.
- Plusieurs études comparent l'efficacité d'une autre molécule à l'amitriptyline en plus du placebo, l'amitriptyline servant de traitement de référence. L'interprétation du résultat de l'étude peut être impactée par le fait que la mesure de l'efficacité de l'amitriptyline n'est pas l'objectif principal de l'étude.
- Toutes les études incluses sont anciennes, la plus récente datant de 2012. La rédaction des articles peut être parfois succincte avec peu de détails méthodologiques rapportés.
- Les études incluses ont toutes de petits effectifs et sont donc de faible puissance. L'effectif le plus important est celui de l'étude de Carette 1994 avec au total 208 patients inclus dans 3 bras.
- Nous avons choisi de sélectionner uniquement les articles en français ou en anglais. De ce fait un article dans une autre langue étrangère a été exclu. Cela peut potentiellement biaiser les résultats de notre étude.

- Nos conclusions ne peuvent pas s'appliquer à une population pédiatrique car nos études incluent uniquement des adultes.
- Bien que nous n'ayons pas trouvé d'article supplémentaire complet sur la plateforme ClinicalTrials.gov, un biais de publication ne peut être exclu de manière empirique.

4. Implication pour la pratique :

La physiopathologie de la fibromyalgie est mal connue, et il n'existe aucun traitement étiologique. De ce fait, aucun médicament n'a l'AMM en Europe, et le traitement de la pathologie est purement symptomatique. Les guides européens de recommandation des bonnes pratiques, rédigés à partir des recommandations de l'EULAR (9), insistent principalement sur la prise en charge multidisciplinaire des patients, avec en priorité les techniques non médicamenteuses telles que l'activité physique. La place des traitements pharmacologiques est jugée secondaire, avec seulement 5 molécules préconisées : l'amitriptyline, la duloxétine, le milnacipran, la prégabaline et le tramadol, dont le bénéfice attendu est comparable et estimé faible.

La méta-analyse de Häuser 2011 conclut, grâce à une comparaison indirecte, à une supériorité de l'amitriptyline sur la duloxétine et le milnacipran, dans le traitement des symptômes de la fibromyalgie, tout en soulignant que le niveau de preuve de l'amitriptyline est inférieur à celui des deux autres traitements. Notre thèse suggère, malgré le manque de preuve, un bénéfice d'un traitement par l'amitriptyline à des doses de 25 à 50 mg, principalement attendu sur les symptômes « troubles du sommeil » avec un effet estimé modéré, mais aussi dans une moindre mesure pour la « douleur », et « l'amélioration de la santé du patient ». De plus, bien que présentant des effets indésirables fréquents mais non graves, notre étude suggère une meilleure adhésion des patients au traitement actif qu'au placebo. Par ailleurs, l'efficacité au cours du temps est mal établie, car mesurée par une seule étude (Carette 1994) durant 6 mois avec une diminution progressive de son efficacité.

La prise en charge d'un patient atteint de fibromyalgie est une prise en charge complexe, multidisciplinaire, souvent insatisfaisante pour le patient et pour le clinicien. En ayant conscience du manque de preuve de son efficacité, et dans le cadre d'une décision médicale

partagée permettant au patient de faire un choix éclairé, et avec des réévaluations régulières de la balance bénéfice – risque au cours du traitement, l’amitriptyline reste une option thérapeutique pour le soulagement des symptômes.

5. Implication pour la recherche :

Notre méta-analyse met en évidence un manque de preuve d’efficacité de l’amitriptyline, et ce malgré des résultats concordants entre presque toutes les études sur le sujet, qui montrent son bénéfice. Ce constat est aussi réalisé par l’étude Moore 2015 qui trouve une amélioration de la douleur d’au moins 50% de son intensité initiale pour un quart des patients traités, mais qui conclut à un niveau de preuve très faible de son résultat et une probable surévaluation de ce dernier. Cette étude a d’ailleurs marqué un tournant décisif dans la place de l’amitriptyline dans le traitement de la fibromyalgie. Ainsi, la molécule a été estimée comme le traitement pharmacologique de référence aux Etats-Unis (36) et en Europe (37) depuis les années 2000 et jusqu’en 2015. Elle a été le gold standard pour tester l’efficacité d’autres molécules sur les symptômes de la pathologie, comme par exemple dans l’étude Zanette 2014 (31). Actuellement son AMM a été retirée aux Etats-Unis du fait de son manque de preuve, au profit d’autres molécules : la duloxétine, la milnacipran et la prégabaline. En Europe et au Canada, les recommandations insistent principalement sur la prise en charge multidisciplinaire des patients, avec en priorité les approches thérapeutiques non médicamenteuses telles que l’activité physique. Ils estiment l’approche pharmacologique comme étant secondaire et l’utilité de ces 4 molécules comme faible.

Cette opposition entre l’efficacité suggérée par les résultats d’une part, et les faiblesses méthodologiques des études incluses qui conduisent à considérer leurs résultats comme non concluants d’autre part, ne pourra être résolue que par la réalisation d’un nouvel essai clinique de bonne qualité.

Concrètement, nous constatons que toutes les études incluses dans notre travail de recherche ont un risque global de biais élevé, avec souvent plusieurs domaines de jugement à risque important.

Il faut noter que la majorité des études sont anciennes et détaillent peu certaines parties de leur méthodologie. Ce manque de précision conduit à l'estimation de risques de biais. C'est particulièrement le cas du domaine évaluant le processus de randomisation. De plus, ces études n'avaient pas les mêmes exigences ou moyens pour garantir une absence de biais résultant de la sélection des résultats rapportés (pas de base d'enregistrement des protocoles). La prise en compte de l'inflation du risque alpha comme biais semble aussi récente, et n'est d'ailleurs incluse ni dans les domaines de biais mesurés par l'outil RoB 2, ni dans la méthodologie d'évaluation du niveau de preuve de Cochrane : GRADE, alors qu'elle l'est dans la méthode plus récente : REB. L'ensemble des biais dans les domaines de jugement sus-cités sont inhérents à la qualité méthodologique intrinsèque du protocole de l'étude. Il serait donc étonnant de retrouver ces biais si des études de meilleures qualités étaient réalisées.

Les deux principales sources de biais extrinsèques au protocole d'étude identifiées sont :

- ❖ Le pourcentage de perdus de vue qui impacte le domaine des données manquantes et le domaine aux écarts aux interventions prévues lorsque les auteurs réalisent une analyse de leurs données en intention de traiter modifiée.
- ❖ Les effets indésirables de l'amitriptyline qui sont spécifiques et fréquents. Ils impactent le domaine des écarts aux interventions (levée du double aveugle) et le domaine de la mesure des critères de jugement (car ces derniers sont subjectifs).

Cependant le contrôle et la minimisation de l'impact de ces biais est aussi réalisable. Nous préconisons que les auteurs utilisent une méthode appropriée de gestion des données manquantes (qui ne soit pas la méthode de la dernière observation reportée, jugée invalide selon la méthodologie REB) et une analyse en intention de traiter (non modifiée). Nous préconisons aussi dans le cadre des essais comparant une molécule active ayant des effets indésirables spécifiques et fréquents, de mesurer l'effet d'une autre molécule active ayant des effets indésirables proches, afin de maintenir le double aveugle. Ainsi, par exemple pour la fibromyalgie, les autres molécules recommandées en traitement de fond pour soulager les symptômes sont la duloxétine, la milnacipran et la prégabaline. Il serait préférable de mesurer sur le même protocole de recherche l'évaluation de l'efficacité contre placebo de l'amitriptyline et de la prégabaline, car ces deux molécules ont des effets indésirables spécifiques proches.

Par ailleurs, la confrontation de notre étude à la littérature suggère un changement important dans l'analyse du niveau de risque de biais par rapport aux méta-analyses plus anciennes. Il

serait pertinent de réaliser une méta-analyse selon la méthodologie REB pour chacune des autres molécules préconisées dans le traitement de la fibromyalgie par l'EULAR (société savante qui construit les guides de bonne pratique médicale en Europe sur la rhumatologie), afin d'avoir une estimation plus juste de leur efficacité.

VI- CONCLUSION

La fibromyalgie est un syndrome complexe, fréquent, handicapant, et responsable d'un coût de santé publique important. Sa physiopathologie est encore insuffisamment connue, bien qu'un dérèglement du système central de la douleur soit l'hypothèse privilégiée. Les troubles du sommeil pourraient être un facteur de risque de développement de la maladie (38). La prise en charge du patient est complexe, multidisciplinaire, et le traitement pharmacologique n'a qu'un rôle secondaire, du fait du faible effet sur les symptômes : douleur, fatigue, troubles du sommeil, diminution de la qualité de vie. L'amitriptyline qui est une molécule active sur le système nerveux central, est l'un des premiers traitements à avoir montré une efficacité sur les deux symptômes principaux : la douleur et les troubles du sommeil. Son bénéfice semble cependant avoir été longtemps surévalué.

Notre méta-analyse confirme le manque de preuves quant à l'efficacité réelle de l'amitriptyline, du fait du haut niveau de risque de biais de l'ensemble des études qui la composent. Elle suggère cependant un bénéfice du traitement, principalement sur les troubles du sommeil (effet estimé modéré), et dans une moindre mesure, dans la réduction de la douleur et l'amélioration globale du patient. Etant donné l'effet limité des traitements, les autres prises en charges doivent être privilégiées. Cependant, la tendance à une adhésion plus importante dans le groupe amitriptyline que dans le groupe placebo, malgré les effets indésirables, laisse à penser que la molécule a toujours sa place dans la pratique, si le traitement est décidé avec l'accord éclairé du patient et que sa balance bénéfice-risque est régulièrement réévaluée. Etant donné la faiblesse méthodologique des études comparant l'amitriptyline au placebo, il paraît d'une part qu'utiliser la molécule comme gold standard dans les essais cliniques est une erreur, et d'autre part qu'une nouvelle étude de bonne qualité méthodologique comparant l'amitriptyline la prégabaline et le placebo est nécessaire.

Il serait aussi pertinent de réaliser une méta-analyse selon la méthodologie REB sur les autres molécules préconisées par l'EULAR dans le traitement symptomatique des patients atteints de fibromyalgie.

VII- FINANCEMENT ET DÉCLARATION DE CONFLITS D'INTÉRÊTS

Les auteures déclarent ne pas avoir reçu de financement pour la réalisation de cette revue systématique de la littérature et ne pas avoir de conflit d'intérêt.

VIII- BIBLIOGRAPHIE

1. Belgrand, L., So, A., Critères de diagnostic de la fibromyalgie, Rev Med Suisse, 2011/286 (Vol.-3), p. 604-608.
2. Sarzi-Puttini P, Giorgi V, Marotto D, Atzeni F. Fibromyalgia: an update on clinical characteristics, aetiopathogenesis and treatment. Nat Rev Rheumatol. 2020;16(11):645-660
3. Allaz, A., Piguet, V., La fibromyalgie : entre certitudes et interrogations, Rev Med Suisse, 2007/116 (Vol.-7), p. 1555-1555
4. Gracely, R.H., Petzke, F., Wolf, J.M. and Clauw, D.J. (2002), Functional magnetic resonance imaging evidence of augmented pain processing in fibromyalgia. Arthritis & Rheumatism, 46: 1333-1343.
5. Moore RA, Derry S, Aldington D, Cole P, Wiffen PJ. Amitriptyline for fibromyalgia in adults. Cochrane Database of Systematic Reviews 2015, Issue 7. Art. No.: CD011824.
6. Ministère de la Santé et des Services sociaux. Algorithme de prise en charge de la fibromyalgie.pdf [Internet]. Disponible sur: https://publications.msss.gouv.qc.ca/msss/fichiers/2021/21-947-01W_fibromyalgie.pdf
7. Amin Thomas.S, Knight. L, Balian. A. Treatment of Fibromyalgia Pain. US Pharm. 2016;41(3):51-54
8. DosquetP, Syndrome fibromyalgique de l'adulte - Rapport d'orientation. pdf [Internet]. Disponible sur: https://www.has-sante.fr/upload/docs/application/pdf/2010-10/syndrome_fibromyalgique_de_ladulte_-_rapport_dorientation.pdf
9. Macfarlane GJ, Kronisch C, Dean LE, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis. 2017;76(2):318-328.
10. Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews BMJ 2021; 372 :n71 doi:10.1136/bmj.n71

11. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; 366: l4898
12. Amitriptyline : substance active à effet thérapeutique [Internet]. VIDAL. Disponible sur: <https://www.vidal.fr/medicaments/substances/amitriptyline-6678.html>.
13. LAROXYL 25 mg, comprimé pelliculé. Résumé des caractéristiques du produit. Dans : Base de données publique des médicaments [Internet]. [cité 11 févr 2022]. Disponible sur : <https://base-donnees-publique.medicaments.gouv.fr/affichageDoc.php?specid=65892928&typedoc=R>.
14. Bousageon R, Blanchard C, Menini T, Pereira B, Naudet F, Kassai B, et al. Rebuild the Evidence Base (REB) project: a method for reviews and meta-analyses focusing on benefit-risk assessment of interest to patients and general practitioners.
15. Schünemann H, Brożek J, Guyatt G, Oxman. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. handbook [Internet]. Disponible sur: <https://gdt.gradepro.org/app/handbook/handbook.html>

Articles inclus dans notre étude:

16. S. Carette, McCain GA, Bell DA, Fam AG, Evaluation of amitriptyline in primary fibrositis. A double-blind, placebo-controlled study. *Arthritis and rheumatism*. 1986;29(5):655-65
17. Don L. Goldenberg, Felson DT, Dinerman H A randomized, controlled trial of amitriptyline and naproxen in the treatment of patients with fibromyalgia. *Arthritis and rheumatism*. 1986;29(11):1371-1377
18. Scudds RA, McCain GA, Rollman GB, Harth M. Improvements in pain responsiveness in patients with fibrositis after successful treatment with amitriptyline. *J Rheumatol Suppl*;19:98-103.
19. Goldenberg DL, Adachi J, Guyatt G, Keller J, Wong B Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials. *Journal of rheumatology*. 1991;18(3):447-451
20. Carette S, Bell MJ, Reynolds WJ, Haraoui B, McCain GA, Bykerk VP, et al. Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. A randomized, double-blind clinical trial. *Arthritis Rheum*. 1994;37(1):32-40.
21. Kempnaers C, Simenon G, Vander Elst M, Fransolet L, Mingard P, de Maertelaer V, Appelboom T, Mendlewicz J. Effect of an antidiemphalon immune serum on pain and sleep in primary fibromyalgia. *Neuropsychobiology*. 1994; 30(2-3):66-72
22. Carette S, Oakson G, Guimont C, Steriade M. Sleep electroencephalography and the clinical response to amitriptyline in patients with fibromyalgia. *Arthritis Rheum*. sept 1995;38(9):1211-7.

23. Goldenberg D, Mayskiy M, Mossey C, Ruthazer R, Schmid C. A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia. *Arthritis Rheum.* nov 1996;39(11):1852-9.
24. Ginsberg F, Mancaux A, Joos E, Vanhove P, Famaey J-P, A randomized placebo-controlled trial of sustained-release amitriptyline in primary fibromyalgia. *Journal of musculoskeletal pain.* 1996;4(3):37-47
25. Hannonen P, Malminiemi K, Yli-Kerttula U, Isomeri R, Roponen P. A randomized, double-blind, placebo-controlled study of moclobemide and amitriptyline in the treatment of fibromyalgia in females without psychiatric disorder. *Br J Rheumatol.* 1998;37(12):1279-1286.
26. Heymann RE, Helfenstein M, Feldman D. A double-blind, randomized, controlled study of amitriptyline, nortriptyline and placebo in patients with fibromyalgia. An analysis of outcome measures. *Clin. Exp. Rheumatol.* 2001;19(6):697-702
27. Fors EA, Sexton H, Götestam KG. The effect of guided imagery and amitriptyline on daily fibromyalgia pain: a prospective, randomized, controlled trial. *J Psychiatr Res.* 2002;36(3): 179-187
28. Braz AS, Morais LC, Paula AP, Diniz MF, Almeida RN. Effects of Panax ginseng extract in patients with fibromyalgia: a 12-week, randomized, double-blind, placebo-controlled trial, *Revista brasileira de psiquiatria (Sao Paulo, Brazil : 1999).* 2013;35(1):21-28

Autres articles:

29. Häuser W, Petzke F, Üçeyler N, Sommer C. Comparative efficacy and acceptability of amitriptyline, duloxetine and milnacipran in fibromyalgia syndrome : a systematic review with meta-analysis. *Rheumatology(Oxford)* 2011;50(3):532-43.
30. Nishishinya B, Urrútia G, Walitt B, et al. Amitriptyline in the treatment of fibromyalgia: a systematic review of its efficacy, *Rheumatology*, 2008, vol. 47 (pg. 1741-6)
31. Zanette SA, Vercelino R, Laste G, Rozisky JR, Schwertner A, Machado CB, et al. Melatonin analgesia is associated with improvement of the descending endogenous pain-modulating system in fibromyalgia: a phase II, randomized, double-dummy, controlled trial. *BMC Pharmacology and Toxicology* 2014;15:40.
32. Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, 273(5), 408–412.
33. Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Controlled clinical trials*, 17(1), 1–12.
34. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org
35. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org
36. Goldenberg, D.L., Burckhardt, C. and Crofford, L. (2004) Management of Fibromyalgia Syndrome. *Journal of the American Medical Association*, 292, 2388-2395.
37. Carville, S. F., Arendt-Nielsen, L., Bliddal, H., Blotman, F., Branco, J. C., Buskila, D., Da Silva, J. A., Danneskiold-Samsøe, B., Dincer, F., Henriksson, C., Henriksson, K.

G., Kosek, E., Longley, K., McCarthy, G. M., Perrot, S., Puszczewicz, M., Sarzi-Puttini, P., Silman, A., Späth, M., Choy, E. H., ... EULAR (2008). EULAR evidence-based recommendations for the management of fibromyalgia syndrome. *Annals of the rheumatic diseases*, 67(4), 536–541.

38. Moldofsky H, Franklin AL, Mously C, Roth-Schechter B, Reynolds WJ. The effect of Zolpidem in patients with fibromyalgia: A dose ranging, double blind, placebo controlled modified crossover study. *J Rheumatol* 1996; 23:529–3

IX- ANNEXES :

1. S. Carette 1986: Evaluation of amitriptyline in primary fibrositis. A double-blind, placebo-controlled study

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles.
Participants	N = 70, femmes et hommes, non hospitalisés, présentant les critères de la fibromyalgie selon les critères proposés par Smythe : - Douleurs généralisées depuis plus de 3 mois - Sensibilité localisée à 12 des 14 sites spécifiques - Sommeil perturbé avec fatigue et raideur matinales - Absence de traumatisme, d'affection neurologique, musculaire, infectieuse, osseuse, endocrine ou d'autre affection rhumatologique - Une vitesse de sédimentation érythrocytaire normale, un taux de créatine phosphokinase normal, un test d'agglutination au latex, des anticorps anti nucléaires et une TSH normaux.
Interventions	Les participants ont reçu soit : - de l'Amitriptyline, initialement à la dose de 10 mg au coucher la première semaine, puis 25 mg à partir de la seconde jusqu'à la quatrième semaine et enfin 50 mg à partir de la cinquième semaine jusqu'à la fin de l'essai. Soit - un Placebo.

Outcomes	Critère de jugement principal : “amélioration significative selon l'évaluation globale du patient : modérée ou marquée ou augmentation du TMS de 8 kg ou plus” Critères de jugement secondaire : Evaluation faite par le patient : - Durée de la raideur matinale en minutes. - Evaluation de la douleur globale sur une échelle analogique de 1 à 10 avec 10 = tolérable. On a également demandé aux patients d'indiquer l'étendue des douleurs ressenties la semaine passée. - Evaluation de la qualité du sommeil comparée aux valeurs initiales, notées de 1 à 3 avec 3 = pas de changement. - Evaluation globale de la maladie comparée aux valeurs initiales sur une échelle de 1 à 5 avec 1 = pire et 5 = amélioration marquée.
----------	---

	Evaluation faite par le clinicien : - Mesure des points sensibles de la fibromyalgie à l'aide d'un dolorimètre de 9 kg. Les scores individuels ont été ajoutés pour donner le " Total myalgic score" (TMS). - Evaluation globale de la maladie comparée aux valeurs initiales sur une échelles de 1 à 5 avec 1 = pire et 5 = amélioration marquée. Autres : - Les effets indésirables
Notes	Critères d'exclusion : Les AINS, les hypnotiques et les antidépresseurs devaient avoir été interrompus au moins 3 semaines avant l'entrée dans l'étude. Seul le paracétamol était autorisé durant l'étude, et chaque dose était enregistrée. Les patients traités avec de l'amitriptyline dans l'année précédente et ceux présentant un antécédent d'hypersensibilité à ce médicament étaient exclus. Les patients avec un antécédent de glaucome, de rétention urinaire, de maladie cardiaque ischémique, d'arythmie cardiaque, ou d'insuffisance cardiaque étaient également exclus. Étude multicentrique : 3 centres universitaires au Canada Durée 9 semaines Analyse en per protocole N = 59 patients ont été finalement analysés : N =27 dans le groupe Amitriptyline et N = 32 dans le groupe Placebo. 11 participants ont abandonné l'étude (7 dans le groupe Amitriptyline et 4 dans le groupe Placebo). Résultat significatif pour l'Amitriptyline versus placebo dans l'évaluation du sommeil avec une amélioration significative à la 5-ème semaine et à la 9-ème semaine. Résultat significatif dans l'évaluation globale de la maladie par le patient avec amélioration significative du groupe Amitriptyline versus Placebo à la 5ème semaine et à la 9-ème semaine.

b) Tableau RoB2

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	Etude multicentrique réalisée dans 3 centres universitaires canadiens. Etude réalisée par le même auteur en 1994 ayant utilisé une table numérique de randomisation. Nous considérons donc qu'il est probable que la séquence pour cette étude soit aléatoire.	PY
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	Voir question 1.1 : Nous considérons qu'il est probable que la séquence d'allocation ait été dissimulées convenablement.	PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	<i>"The randomisation resulted in 2 comparable groups with respect to age, sex, duration of morning stiffness, pain analog scores, and total myalgic score The duration of symptoms before entry study was slightly longer in patients receiving placebo P = 0.04"</i>	N
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions envisagées (effet de l'adhésion à l'intervention, per protocole)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Amitriptyline capsule identical to placebo Effets indésirables : « side effects were reported by 19 patients (70%) receiving amitriptyline who completed the study and 4 (12%) taking placebo. These side effects were, for the most part, drowsiness and xerostomia. »	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 Les interventions non protocolaires	Pas d'interventions non protocolaires	Non concerné

importantes étaient-elles équilibrées entre les groupes d'intervention ?		
2.4 [Le cas échéant] Y a-t-il eu des échecs dans la mise en œuvre de l'intervention qui auraient pu affecter les résultats ?		NI
2.5 [Le cas échéant] Y a-t-il eu un non-respect du schéma d'intervention assigné qui aurait pu affecter le résultat des participants ?	- 4 exclusions pour inobservance « lack of cooperation » (3A, 1P) - 4 arrêts pour EI « drowsiness (1 A, 1P), agitation (1 A), gastrointestinal symptoms (1 P) - 1 pour insuffisance d'effets thérapeutique (1 A) Les 2 arrêts de traitements pour pathologie intercurrente sont donc justifiés (donc non retirés) $N = 4 + 4 + 1 \Rightarrow 9/70 \Rightarrow 12.8\%$ de mauvaise observance au traitement (11/70 = 15.7% si on compte tout)	PY
2.6 <u>Si N / PN / NI à 2.3 ou Y / PY / NI à 2.4 ou 2.5</u> : Une analyse appropriée a-t-elle été utilisée pour estimer l'effet du respect de l'intervention ?		PN
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	Données manquantes pour 11/70 patients randomisés soit 15.7% des patients	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas d'analyse de sensibilité ou de méthode de correction de biais	PN
3.3 <u>Si N / PN à 3.2</u> Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	11 / 70 = 15.7% de données manquantes dont 9/70 = 12.8% en lien avec vraie valeur : - lack of cooperation 4 - side effects : 4 - insuffisance thérapeutique : 1	PY

3.4 <u>Si Y/ PY / NI à 3.3</u> : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	Presque 2 fois plus d'abandon dans le groupe amitriptyline que placebo (7 amitriptyline et 4 placebo) et cause de perdu de vue en lien avec la vraie valeur.	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 <u>Si N / PN / NI à 4.1 ou 4.2</u> : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Doubles aveugles non maintenus	PY
4.4 <u>Si Y / PY / NI à 4.3</u> : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critères subjectifs	PY
4.5 <u>Si Y / PY / NI à 4.4</u> : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole.	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.	Une seule analyse pour chaque critère de jugement	PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Nous considérons que la randomisation et la dissimulation de la séquence d'allocation aient été réalisées dans les règles compte tenu de l'équipe d'investigation Données démographiques de la population initiale ne suggérant pas de problème de randomisation
Biais dû aux écarts à l'intervention reçue: effet de l'adhésion à l'intervention	Risque élevé	Double aveugle a priori non maintenu (Effets indésirables spécifiques de l'amitriptyline) Manque d'adhésion à l'intervention Pas d'analyse appropriée pour estimer l'effet de l'adhésion
Biais dû aux données manquantes	Risque élevé	17,5% de données manquantes Absence d'information concernant la gestion des données manquantes Données manquantes en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Double aveugle à priori non maintenu Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement
Autres biais		Pas de méthode d'ajustement du risque alpha rapportée Pas d'informations sur les conflits d'intérêts
Risque de biais global	Risque élevé	

d) Tableaux de résultats :

Auteur : Carette année : 1986						
Étude randomisée en groupes parallèles 2 bras						
Amitriptyline vs placebo						
Analyse en per protocole						
Patients ayant fini l'étude N = 59, amitriptyline n = 27, placebo n = 32						
	Critère de jugement	Mesure du critère	Valeur rapportée	Amytriptyline Résultats à la 9 ^{ème} semaine	Comparaison avec placebo à la 9 ^{ème} S	Significatif ou non vs placebo
Critère de jugement principal	Amélioration modérée ou marquée globale des symptômes de la fibromyalgie (Par le patient)	Echelle de 1 à 5 1 = pire 2 = inchangé 3 = amélioration minime 4 = amélioration modérée 5- amélioration marquée	Population améliorée en pourcentage (cotent 4 ou 5 sur 5)	A : 63% (17 patients sur 27) Modéré : 11 Marqué : 6 A la 5 ^{ème} semaine : Modéré 11/27 Marqué: 4/27	P : 32 % (10 /32 patients) Modéré : 5 Marqué : 5 A la 5 ^{ème} S : Modéré :4/32 Marqué:3/32	P = 0,003 à la 5 ^{ème} semaine p = 0.002
	Points sensibles de la fibromyalgie	"Total myalgic score"	Somme des points douloureux de chaque patient par groupe	25.2 ± 8.1	26.2 ± 8.9	Non significatif (p non donné)
Critère de jugement secondaire	Evaluation globale des symptômes de la fibromyalgie (Par le patient)	Echelle de 1 à 5 1 = pire 2 = inchangé 3 = amélioration minime 4 = amélioration modérée 5- amélioration marquée	Population améliorée en pourcentage (cotent 4 ou 5 sur 5)	70% (19 patients sur 27) Minime : 2 Modéré : 11 Marqué : 6 A la 5 ^{ème} S: 77 % (21/27 patients)	50% (16 / 32 patients) Minime : 6 Modéré : 5 Marqué : 5 5 ^{ème} S : 43% (14/32 patients)	p = 0.12 non significatif à la 5 ^{ème} S p = 0.008
Critère de jugement secondaire	Raideur matinale (patient)	Durée	Minutes	48 ± 61	66 ± 76	p = 0.12
	Douleur (patient)	EVA	Moyenne +/- SD	4.3 ± 3.0	5.0 ± 3.0	p = 0.12
	Qualité du sommeil (patient)	1 = plus reposé 2 = moins reposé 3 = aucun changement	Population améliorée en pourcentage	70% des patient améliorés (à 5 semaines 70%)	40% des patients améliorés (5°S : 34%)	p = 0.02 (5°S, p = 0.008)
	Evaluation globale des symptômes de la fibromyalgie (médecin)	Echelle de 1 à 5 1 = pire 2 = inchangé 3 = amélioration minime 4 = amélioration modérée 5- amélioration marquée		n = 18/27 Minime : 3 Modérée : 8 Marquée : 7 5 ^{ème} semaine : n = 19/27 Minime : 8 Modérée : 6 Marquée : 5	n = 14/27 Minime : 6 Modérée :6 Marquée : 2 5 ^{ème} S n =12/27 Minime : 7 Modérée : 2 Marquée : 3	Pas de comparaison entre les groupes amitriptyline

2. Don. L. Goldenberg 1986: A randomized, controlled trial of amitriptyline and naproxen in the treatment of patients with fibromyalgia

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles.
Participants	N = 62 Adultes âgés de 21 à 69 ans Les participants devaient répondre aux critères cliniques de la fibromyalgie modifiés par Yunus et al : - Des courbatures et douleurs généralisées ou raideur proéminente impliquant 3 sites anatomiques ou plus pendant au moins 3 mois ; - Une absence de causes sous-jacentes ou de traumatismes répétitifs ou de maladie systémique ; - Au moins 6 points sensibles de la fibromyalgie typiques, modérément ou sévèrement sensibles ; - Au moins 3 des éléments suivants : une modulation des symptômes par l'activité, la météo, l'anxiété ou le stress, la présence d'un mauvais sommeil, une fatigue générale ou une asthénie, des maux de tête chroniques, une irritabilité, un syndrome intestinal, ou un gonflement et un engourdissement subjectif. Par ailleurs, les patients devaient obtenir un score de 4 ou plus sur au moins une des deux échelles visuelles analogiques autoadministrées de 10 mm évaluant les symptômes ou la douleur de la fibromyalgie.
Interventions	Les participants ont reçu soit : - Du Naproxène à la dose de 500 mg deux fois par jour et de l'Amitriptyline à dose de 25 mg le soir Soit - Du Naproxène à la dose de 500 mg deux fois par jours et un Placebo d'Amitriptyline le soir Soit - De l'Amitriptyline à la dose de 25 mg le soir et un Placebo de Naproxène deux fois par jour. Soit - Un Placebo de Naproxène deux fois par jour et un Placebo d'Amitriptyline le soir.
Outcomes	<u>Critères de jugement principal :</u> Le pourcentage de patient ayant présenté une amélioration modérée ou marquée de leurs symptômes et/ou du TMS (total myalgic score) de 8 kg ou plus

	<u>Critère de jugement secondaire :</u> - Evaluation globale des symptômes par le patient et par l'investigateur sur une EVA de 0 à 10 mm. - Evaluation de la douleur ou de la raideur par le patient sur une EVA de 0 à 10 mm. - Evaluation de la fatigue par le patient sur une EVA de 0 à 10 mm. - Evaluation des troubles du sommeil sur la base de 2 questions avec des réponses notées sur une EVA de 0 à 10 mm : A) Avez-vous des difficultés à dormir ? B) Comment vous sentez vous le matin au réveil ? - Nombre de points sensibles positifs du fibromyalgie noté de 0 à 2 pour chaque zone. Un score total de sensibilité (TMS) est ensuite calculé en additionnant les résultats de chaque zone. - Effets indésirables Observance
Notes	Critères d'exclusion : La présence d'un antécédent d'ulcère gastroduodénal, d'arythmie cardiaque, la prise de médicaments qui ne pouvaient pas être arrêtés. Tous les patients devaient avoir arrêté les antalgiques, les anti-inflammatoires, les antidépresseurs, les somnifères et autres médicaments agissant sur le système nerveux central au moins 72 heures avant leur première visite. Le paracétamol était autorisé en cas de douleur intense. Durée de l'essai : 6 semaines Analyse en ITTm Patients analysés N =58 4 abandons : 1 dans chaque groupe Résultat significatif pour l'Amitriptyline versus placebo - dans l'amélioration du score de points sensibles à semaine 4 ($p < 0,05$). - dans l'amélioration dans la douleur/raideur évaluée par le patient à semaine 4 ($p < 0,01$). - dans l'amélioration de la fatigue à semaine 6 ($p < 0,05$). - dans l'amélioration du sommeil à semaine 2 et semaine 6 ($p < 0,05$). - dans l'amélioration globale de la fibromyalgie évaluée par le patient à semaine 4 et semaine 6 ($p < 0,05$). - dans l'amélioration globale de la fibromyalgie évaluée par l'examineur à semaine 2, semaine 4 et semaine 6 ($p < 0,01$).

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« Randomly assigned to 1 of 4 treatment groups using a method that assured balanced assignment (19) » “(19) Peto R, Pike MC, Armitage P : Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Introduction and design. BR J Cancer 34 :585-612, 1976 “ Dans ce texte cité => « At the start of a trial, each participating centre may be given an ordered set of sealed envelopes. [...]Each envelope contains the instruction as to witch treatment schedule to follow for a particular patient. When a patient who appears to satisfy the admission criteria is found, he is formally admitted, and only then is the envelope opened to discover his treatment. » p 598	Y
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		Y
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	Pas de différence de base significative entre les 4 groupes de traitements pour l'ensemble des critères de base testés (race, durée des symptômes de la fibromyalgie, troubles du sommeil, fatigue matinale, traitements antérieurs par AT ou autre antidépresseur, nombre et score de tender points.	N
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)

Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	« Patients, as well as the examining physician, were « blinded » as to specific treatment. » « 8 patients complained of adverse side effects, but did not discontinue the study medication. These complaints included dry mouth in 4 patients (all in groups 1 and 3) »	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?	Cf ci dessus	PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITTm avec 4 drop outs /62 patients inclus => 6.45% « Patient drop outs were defined as those patients who failed to return for a second visit. All other patients were included in the analysis. [...] there were 4 drop outs, 1 in each treatment group.	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	6.45% de patient exclus post randomisation et analyse non appropriée	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	- 4 drops out exclus de l'analyse - 2 patients qui n'ont pas terminés l'étude dans le groupe 1 mais inclus dans l'analyse Donc : 6/62 => 9.677% données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas d'informations sur la gestion des données manquantes	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Causes de données manquantes : - en lien avec l'état de santé du patient : 2/6 (G1 et 4) - inconnues : 4/6 Répartition des données manquantes selon l'analyse - Par groupe de traitement : G1 : 3, G2 : 1, G3 : 1, G4 : 1 Donc données manquantes de cause inconnues possiblement aussi en lien avec vraie valeur	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?		NI
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Examineur : patient et/ou médecin Double aveugle non maintenu	PY
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critère de jugement subjectif	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Quelques Inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Séquence d'allocation aléatoire Dissimulation de la séquence d'allocation bien réalisée Données démographiques de la population initiale ne suggérant pas de problème de randomisation
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle a priori non maintenu (Effets indésirables spécifiques de l'amitriptyline) Analyse en ITTm non adaptée avec plus de 5% de données manquantes Probable impact sur les résultats
Biais dû aux données manquantes	Risque élevé	9.6 % de données manquantes Absence d'information concernant la gestion des données manquantes Données manquantes en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Double aveugle à priori non maintenu Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement
Autres biais		Pas d'informations sur les conflits d'intérêts Multiplicité des critères de jugements mais ajustement du risque alpha : « t-test adjusted by Duncan's multiple range test for multiple comparisons . A 2-tailed value of 0.05 was considered significant. »
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur: Goldenberg Année: 1986					
Patients inclus : n = 62, patients analysés n = 58 Etude randomisée contre placebo en 4 groupes : amitriptyline + placebo (A+ P), amitriptyline + Naproxène (A+N), naproxène + placebo (B+P), placebo (P)					
Pas de séparation critère de jugements primaires et secondaires					
	Critère de jugement	Mesure du critère	Valeur rapportée	Groupe amitriptyline + placebo à la 6 ^{ème} semaine (S)	Significatif ou non par rapport au placebo à la 6 ^{ème} S
Critère de jugements primaires et secondaire	Evaluation Globale des symptômes de fibromyalgie	Par le patient Echelle Visuelle Analogique (0 -> 10cm)	Moyenne	A : 4.5 P : 7.9	p < 0.05
	Douleur ou raideur			A : 5.4 P : 7.3	p < 0.01
	Fatigue			A : 4.3 P : 7.5	p < 0.05
	Trouble du sommeil			A: 3.0 P: 6.7	p < 0.05
	Sensation au premier réveil			Pas de résultat dans le texte, ni dans les figures	
	Tender point	Tender point score		A: 11.6 Pas de donnée pour placebo (graphique)	p < 0.05
	Evaluation globale par le médecin	Echelle visuelle analogique (0 -> 10 cm)		A: 4.1 P: 8.8	p < 0.01

3. Ra. Scudds 1989: Improvements in pain responsiveness in patients with fibrositis after successful treatment with amitriptyline

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en cross-over avec 2 bras d'intervention.
Participants	N = 69 Femmes et hommes, non hospitalisés, présentant les critères de la fibromyalgie selon les critères proposés par Smythe et Moldofsky: - Des douleurs musculaire généralisées depuis plus de 3 mois - Une sensibilité localisée à 12 des 14 sites spécifiques - Un sommeil non réparateur - Une fatigue et une raideur matinale - Une vitesse de sédimentation érythrocytaire normale, un taux de TSH normal ainsi que des radiographies normales.
Interventions	Les participants ont reçu soit : - De l'Amitriptyline au coucher à la dose de 10 mg la première semaine, puis 25 mg la seconde et enfin à la dose de 50 mg les deux dernières semaines Soit - Un Placebo au coucher
Outcomes	<u>Critère de jugement primaire :</u> “La douleur et la réactivité à la douleur sont-elles influencées de manière prévisible par un traitement par amitriptyline à faible dose. » En estimant que la réponse est positive si “amélioration cliniquement significative du score myalgique total de 50 %” <u>Autres critères mesurés :</u> - Score myalgique total calculé comme la somme des mesures des points sensibles de la fibromyalgie à l'aide d'un dolorimètre - Seuil de la douleur défini comme la somme des seuils de douleurs de 4 points considérés comme “ non sensibles” - Tolérance à la douleur- Intensité de la douleur (Mc Gill Pain Questionary)- Évaluation globale par le patient de l'efficacité globale du traitement (échelle ordinale à 5 points)
Notes	Critères d'exclusion : Les AINS, les hypnotiques et les antidépresseurs devaient avoir été interrompus au moins 3 semaines avant l'entrée dans l'étude. Seul le paracétamol était autorisé durant l'étude, et chaque dose était enregistrée. Les patients traités avec de l'amitriptyline dans l'année précédente et ceux présentant un antécédent d'hypersensibilité à ce médicament étaient exclus. Les patients avec un antécédent de

	glaucome, de rétention urinaire, de maladie cardiaque ischémique, d'arythmie cardiaque, ou d'insuffisance cardiaque étaient également exclus. Etude monocentrique Durée totale : 10 semaines Période traitement : 4 semaines Période de wash-out : 2 semaines Analyse en ITTm Participants analysés N = 36 Perdus de vue N =3 Résultat significatif pour l'amitriptyline versus Placebo pour l'amélioration globale du traitement (p<0,01)
--	--

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« A completely randomized double blind crossover design was employed. Patients were randomly assigned to one of the 2 groups »	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« Using MANOVA, no significant effect was found for either group [...] or when grouped by time » cependant ce test est utilisé sur la population finale (36 patients) ayants fini l'études et non sur les 39 patients randomisés.	NI
Evaluation du risque de biais		Quelques inquiétudes

Biais dû à l'effet « carry-over »		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	A priori ratio 1 :1 : absence d'information sur le ratio utilisé mais après exclusion de 3 patients pour effets indésirables on a groupe 1 = 19 patients au final avec 1 perdu de vue soit 20 patients initialement ; et groupe 2 = 17 patients au final avec 2 patients perdus de vue soit 19 patients initialement.	PY
S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?		Non concerné
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	Wash-out période de 2 semaines, puis 4 semaines de traitement soit 6 semaines avant mesure du critère de jugement principal sur la deuxième période. Donc effet carry-over peu probable.	PY
Evaluation du biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	« The amitriptyline was in capsules which were identical to the placebo capsule » « Both subjects and tester were blinded as to treatment group for the duration of the trial. » Pas d'information concernant les effets indésirables spécifiques de l'amitriptyline.	NI
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		NI
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné

2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITTm avec 7.7 % de données manquantes (3 abandons sur 39 patients)	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	Analyse inappropriée avec 7,7% de données manquantes	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	3 perdus de vue sur 39 participants soit 7.7 % de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas d'information concernant la gestion des données manquantes	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Abandons pour effets indésirables (2 patients) et manque d'efficacité (1 patient)	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	Plus de 5% de données manquantes et abandons en lien avec la vraie valeur	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?	“Pain responsiveness was assessed using a 9 kg dolorimeter“	PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?	Non même méthodologie de mesure « Pain responsiveness was assessed using a 9 kg dolorimeter » « One tester took all measurements in the trial. » « all evaluations took place in the same environment between the hours of 8 am and 6 pm. »	PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils	Doute sur le maintien du double aveugle	NI

au courant de l'intervention reçue par les participants à l'étude ?		
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critère de jugement subjectif Evaluateur : patient	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information sur un enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	“other measures, such as hypochondriasis, depression, anxiety, sickness impact and daily pain diary were also taken during the study. However, these will not be reported here.”	PY
5.3 ... Plusieurs analyses éligibles des données.	3 analyses réalisées et toutes présentées : Analyse 1 = amitriptyline first group placebo first group Analyse 2 = 2 groups based on the treatment periods Analyse 3 = « improved group » and « unchanged group »	PN
5.4 : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Risque élevé

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Peu d'informations sur la randomisation Absence de données concernant la population de base
Biais dû à l'effet carry-over	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Pas assez d'informations pour statuer concernant le maintien du double aveugle Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	7.7 % de données manquantes Absence d'information concernant la gestion des données manquantes Données manquantes a priori en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Maintien du double aveugle non statué Evaluateur : patient Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Risque élevé	Pas d'information sur les intentions d'analyse préalables à l'essai. Protocole a priori non enregistré préalablement Sélection des résultats mesurés
Autres biais		Absence d'informations concernant les conflits d'intérêts Méthode d'ajustement du risque alpha par la méthode de Tukey
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Scudds Année : 1989

Crossover randomisé en double aveugle contre placebo

Molécules : placebo (P), amitriptyline (A)

39 patients randomisés, 36 patients analysés

Analyse en ITTm

Abréviations : Pas d'Information (PI) placebo (P), amitriptyline (A), Score Myalgique Total (SMT)

	Critère de jugement	Mesure du critère	Valeur rapportée	Résultats		Significatif ou non	
Critère de jugement primaire	Douleur	Amélioration > 50% SMT		Pas de valeur rapportée		PI	
Critère de jugement secondaire		Seuil de douleur	Test univarié F pour le temps avec (3.102)df	Hypoth SS 152.69, Error SS 2411.00 Hypoth MS 50.89, Error df 28.28 F 1.90		P = 0.135 non significatif	
		Score myalgique total		Hypoth SS 952, Error SS 2836.5 Hypoth MS 317.56, Error df 31.51 F 10.07		P < 0.001 En faveur de A	
			Post hoc Tukey's contrasts	HSD = 3.74		P < 0.05 En faveur de A	
			Post hoc contrasts	A : 14.32kg P : 10.51 kg	Amélioration score 36% avec A	P < 0.05 En faveur de A	
		McGill Pain questionnaire	Test univarié F pour le temps avec (3.102)df	Hypoth SS 354.06, Error SS 2511.03 Hypoth MS 118.02, Error df 27.90 F 4.23		p = 0.008	
		Notation de la douleur		Post hoc Tukey's contrasts	HSD = 3.54		P < 0.05
				Post hoc contrast	A : 3.42 P : 8.42	Amélioration score 64% plus bas avec A	
		Tolérance à la douleur	Somme des niveaux de tolérance à la douleur pour chacun des 4 points décrits pour le seuil de douleur	Test univarié F pour le temps avec (3.102)df	Hypoth SS 276.29, Error SS 8239.23 Hypoth MS 92.09, Error df 91.54 F 1.00		P = 0.394 Non significatif
Efficacité globale du traitement par le patient		Nombre de patients ayant ressenti une amélioration	Echelle ordinale à 5 points	27 patients sous AT et 8 sous placebo X ² = 21.6		P < 0.001	

4. R. Jaeschke 1991: Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, crossover en "single case" ou "N-of1 RCT"
Participants	N = 22 Adultes présentant un diagnostic de fibromyalgie suivis dans un même cabinet de rhumatologie
Interventions	Les participants ont reçu soit : - De l'Amitriptyline Soit - Un Placebo
Outcomes	Critère de jugement principal : "Est-ce que l'amitriptyline apporte quelque chose par rapport au placebo ?" Evaluation du critère de jugement grâce à : Un score composite avec pour chaque item une réponse notée de 0 à 7 incluant - Manque d'Energie - Fatigue - Douleur - Raideur matinale - Troubles du sommeil - Céphalées - Troubles intestinaux Nombre de points sensibles évalués par un rhumatologue Critères secondaires : Le niveau de confiance du médecin dans son plan d'action Question posée : "Dans quelle mesure vous vous sentez à l'aise avec votre plan de traitement" avec une réponse notée de 0 à 7
Notes	Critères d'exclusion : non fournis Durée totale : variable selon les individus minimum « 3 treatment pairs, each consisting of 2 two weeks treatment periods » 3 x (2 semaines placebo + 2 semaines amitriptyline) sans période de wash-out, et ajout d'autant de période de « treatment » (= 2 semaines placebo + 2 semaines amitriptyline) que si jugé nécessaire par le patient / le médecin Période traitement : 2 semaines Période de wash-out : 0

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	Peu d'informations concernant la randomisation « the order of administration within each pair determined by random allocation »	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	Absence d'informations	NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	Absence d'informations données sur la population initiale hormis l'âge	NI
Evaluation du risque de biais		Quelques Inquiétudes

Biais dû à l'effet « carry-over »		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	Chaque patient est son contrôle car étude en "single case" donc ratio de 1 : 1	PY
S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?		
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	« The results of the 3 trials suggested that for fibromyalgia amitriptyline acts quickly after it is started, and ceased to act quickly when discontinued.^{19??} "We used 2-week treatment periods in or subsequent N-of-1 RCT in these patients » « Each patient was asked to rate the severity of 7 symptoms at weekly intervals » Référence bibliographique citée : ^{19??} Guyatt GH, Heyting A, Jeaschke R, Keller J, Adachi JD, Roberts RS, N of 1 randomized trials : The options. Submitted for publication	PY

	« <i>Determining the Rapidity of Onset and Termination of Action : We concluded that the onset and termination of treatment effect was rapid after two N of I RCTs in which an advantage of drug over placebo was evident during the first week of each 4-week treatment period. Although subsequent studies using 2-week periods have confirmed this impression, investigators studying a previously untested drug may wish to have further evidence before shortening treatment periods. »</i> Par ailleurs demi-vie de l'amitriptyline courte.	
Evaluation du domaine		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Absence d'informations concernant les effets secondaires indésirables spécifiques de l'amitriptyline Les patients sélectionnés ont participé à un essai ouvert permettant l'adaptation de la dose d'amitriptyline optimale pour chacun. Ils connaissent donc les effets indésirables de l'amitriptyline qui sont très fréquents	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?	“ One patient underwent 2 N-of-1 RCT, one with 10 mg and the other with 50 mg of the active drug daily “ “Those patients who reported improvement on the initial or subsequent dose were enrolled in an N-of-1 RCT at the dose that gave an apparent benefit in the open trial”	PY
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?	Patient ayant participé deux fois à l'étude: mesures supplémentaires du résultat	PY

2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		NI
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITT	Y
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		Non concerné
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes

Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	Etant donné que l'essai s'arrête dès que le souhaite un participant ou son médecin, il n'y a pas eu de perdu de vue. Par contre les temps de suivi varient entre les patients. Donc si la méthodologie de l'essai est valable alors il n'y a pas de données manquantes.	PY
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?		PY
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?		Non concerné
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?		Non concerné
Evaluation du risque de biais		Risque faible

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats

Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils	A priori double aveugle levé	PY

au courant de l'intervention reçue par les participants à l'étude ?		
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critère de jugement subjectif Evaluateur : patient et/ou médecin	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	1 ^{ère} analyse questionnaire général, puis deuxième par item, résultats présents dans l'étude. Pas de comparaison au placebo.	PY
5.3 ... Plusieurs analyses éligibles des données.		PN
5.4 : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Risque élevé

c) **Résumé des domaines et jugement global :**

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Peu d'informations concernant la randomisation Pas de données concernant la population de base hormis l'âge
Biais dû à l'effet carry-over	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle levé Ecart au protocole Analyse en ITT
Biais dû aux données manquantes	Risque faible	Absence de données manquantes
Biais dû à la mesure de l'outcome	Risque élevé	Levée du double aveugle Evaluateur : patient et/ou médecin Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Risque élevé	Pas d'information sur les intentions d'analyse préalables à l'essai. Protocole a priori non enregistré préalablement Sélection des résultats Les résultats pour les deux périodes sont disponibles
Autres biais		Absence d'informations concernant les conflits d'intérêts Pas de méthode d'ajustement du risque alpha mentionnée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Jaeschke Année : 1991					
	Critère de jugement primaire				
	Différence entre Score composite de chaque période de traitement (c'est-à-dire 2 semaines amitriptyline - 2 semaines de placebo) Manque d'Energie Fatigue Douleur Raideur matinale Troubles du sommeil Céphalées Troubles intestinaux			Différence du nombre de tender points sous Amitriptyline et sous placebo par période de 4 semaines	
Valeur rapportée	Résultats	Effet positif ou négatif de l'amitriptyline sur placebo	Significatif (p < 0.05) ou non (P > 0.05)	Résultats	Significatif (p < 0.05) ou non (P > 0.05)
	Positif (en faveur amitriptyline) Négatif (en faveur placebo)			Positif Négatif	
Patient 1	-0.43 / - 0.64 / 0	Négatif	P = 0.10	2 / 1 / 1	P = 0.03
Patient 2	0.57 / - 2.57 / -1.07		P = 0.19	0 / -8 / -2	P = 0.15
Patient 3	0.08 / 0.43 / -1.07		P = 0.33	7 / 2 / -4	P = 0.33
Patient 3 bis	-0.36 / -0.14 / 0.64		P = 0.45	- 10 / 6 / 1	P = 0.43
Patient 5	1.85 / 0.78 / 0.54	Positif	P = 0.06	4/2/8	P = 0.06
Patient 6	1.35 / 1.36 / 0.79	Positif	P = 0.01	6 / -1 / 1	P = 0.22
Patient 7	- 0.64 / 1.00 / -0.29		P = 0.49	0 / 2 / 2	P = 0.09
Patient 8	0.42/ 0.71 / 0.66	Positif	P = 0.01	6 / 3 / -1	P = 0.16
Patient 9	0.05/- 0.22/0.57/0.36		P = 0.10	1/0/1/-2	P = 0.50
Patient 10	-0.21/1.28/-1.50		P = 0.44	-1/-1/-11	P = 0.16
Patient 11	0.47 / 0.93 / 0.86	Positif	P = 0.02	3	Absent
Patient 12	4.29 / 3.15 / 0.78 / 4.49	Positif	P = 0.02	5 / 10 / 12	P = 0.02
Patient 13	0.68 / -0.25 / -0.09		P = 0.37	0 / -3 / -2	P = 0.10
Patient 14	0.86 / 1.39 / 1.85	Positif	P = 0.02	4/3/0	P = 0.10
Patient 15	0.86 / 1.43 / 0.65 / 1.86	Positif	P = 0.01	3/7/4/5	P < 0.01
Patient 16	0.64/0/0.5	Positif	P = 0.10	-4 / 1 / 4	P = 0.45
Patient 17	- 0.08 / 0.86 / 1.07 / 1.15		P = 0.04	4 / 4 / 11 / 5	P = 0.02
Patient 18	0.64 / 1.08 / -0.36 / 0.79 / - 0.64 / 1.50		P = 0.10	2 / 3 / 0 / 0 / 0 / 1	P = 0.06
Patient 19	0.07 / -0.07 / 0.15		P = 0.26	5 / -1 / 2	P = 0.19
Patient 20	0.29 / -1.42 / 0.14		P = 0.31	0 / -2 / 0	P = 0.21
Patient 21	- 0.29 / 1.08/1.50		P = 0.15	-2 / -2 / 1	P = 0.21
Patient 22	0.35 / - 2.00 / -0.92		P = 0.17	2 / 4 / -9	P = 0.42
Patient 23	1.22 / 1.07 / -0.08 / 0.50		P = 0.06	3 / 0 / 0 / 2	P = 0.10

5. S. Carette 1994 : Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. A randomized, double-blind clinical trial

a) Caractéristiques de l'étude :

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles.
Participants	N = 208 Adultes de plus de 18 ans, remplissant les critères de l'American College of Rheumatology de 1990 pour la classification de la fibromyalgie. De plus, les patients devaient avoir un score supérieur ou égal à 4 sur au moins une des deux EVA auto-administrés de 10 cm (Une évaluant la douleur et une évaluant le bilan global des symptômes de la fibromyalgie). Les participants devaient avoir des résultats normaux concernant le dosage de la TSH et des CPK, ainsi qu'un test de vitesse de sédimentation érythrocytaire normal.
Interventions	Les participants ont reçu soit : - De l'Amitriptyline et un placebo de Cyclobenzaprine. La dose initiale d'Amitriptyline était de 10 mg par jour pendant la première semaine puis de 25 mg par jour pendant les 11 semaines suivante et à partir de la 12-ème semaine la dose était de 50 mg par jour. Soit - De la Cyclobenzaprine et un placebo d'Amitriptyline. La dose initiale de Cyclobenzaprine était 10 mg par jour pendant la première semaine puis de 20 mg par jour pendant les 11 semaines suivante et à partir de la 12-ème semaine la dose était de 30 mg par jour. Soit - Un placebo d'Amitriptyline et un placebo de Cyclobenzaprine administrés aux mêmes horaires que les molécules actives.
Outcomes	<u>Critère de jugement principal :</u> 50% d'amélioration sur au moins 4 des 6 critères suivants : - La douleur sur EVA de 0 à 10 mm - La fatigue sur EVA de 0 à 10 mm - Le sommeil sur EVA de 0 à 10 mm - Évaluation globale de la fibromyalgie par le patient sur une EVA de 0 à 10 mm - Evaluation globale de la fibromyalgie par le médecin examinateur sur une EVA de 0 à 10 mm - - augmentation de 1kg au score des douleurs musculaires, Mesure des points sensibles de la

	Fibromyalgie “ Tender points “par un dolorimètre de 9 kg. Les scores individuels ont été additionnés pour donner le “Total Myalgic score” <u>Critères secondaires :</u> - La sensation au réveil sur EVA de 0 à 10 mm - La raideur matinale sur une EVA de 0 à 10 mm - Evaluation de la douleur par le McGill Pain Questionnaire - Evaluation du handicap fonctionnel par le Sickness Impact Profile (SIP) et le Health Assessment Questionnaire (HAQ) - Le statut psychologique évalué par le Anxiety and depression scales of the Arthrititis Impact Measurement Scales (AIMS) - Evaluation de la personnalité par le Minnesota Multiphasic Personality Inventory (MMPI)
Notes	Critères d'exclusion : Les patients étaient exclus s'ils présentaient des signes de maladie rhumatismale inflammatoire, endocrinienne non traitée, un trouble neurologique, infectieux ou osseux. Les patients précédemment traités par amitriptyline ou cyclobenzaprine ont été également exclus, de même que ceux ayant des antécédents de glaucome, de rétention urinaire, d'infarctus du myocarde, d'arythmie ou d'insuffisance cardiaque congestive. Les anti-inflammatoires non stéroïdiens, hypnotiques et les antidépresseurs ont été arrêtés au moins 3 semaines avant l'entrée dans l'essai. Le seul traitement hors étude autorisé pendant l'essai était le paracétamol. Etude multicentrique Durée de l'étude : 6 mois Analyse en ITTm Participants analysés : 184 à 6mois 52 abandons : 14 dans le groupe Amitriptyline, 24 dans le groupe Cyclobenzaprine et 14 dans le groupe Placebo. Résultat significatif pour l'Amitriptyline versus Placebo pour le critère de jugement principal à 1 mois uniquement (p=0.02).

b) Tableau RoB 2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« The assignment scheme was generated using a table of random numbers. »	Y
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	Etude multicentrique réalisée dans 3 centres universitaires canadiens. Nous considérons qu'il est probable que la séquence d'allocation ait été dissimulée convenablement.	PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« the 3 study groups did not differ significantly with respect to age, sex distribution, education level, or marital status. » « the three groups did not differ significantly with regard to baseline measure of fibromyalgia symptoms, functional scores, or psychological status »	N
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention,)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	« The nature and frequency of side effects did not differ between the 2 active drugs »	PN
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PN
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné

2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	« Two types of analysis were performed : [...] (intention-to-treat) and one based on the results for only those patients who continued taking the study medication through-out of the trail. Since the 2 analyses yielded similar results, only those from the intent-to-treat analysis are reported. » Bien que les auteurs disent avoir réalisé une analyse en ITT, les résultats fournis dans le tableau 4 ont été analysés en ITTm on considère donc que l'analyse réalisée est en ITTm avec plus de 5 % de données manquantes. 184 patients analysés sur 208 à 6 mois = 11.5% de données manquantes.	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	52 patients ayant quittés / ont été exclus de l'étude sur 208 patients => 25% de données manquantes Pourtant dit « Follow up data were obtained on 186 patients after 3 months and 184 patients after 6 months (88%) » soit plus de 11.5% de perdus de vue.	PN
3.2 Existe-il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Absence d'informations concernant la gestion des données manquantes.	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Données manquantes dépendante de la vraie valeur : - abandon pour : - Lack of response, adverse drug reaction - Exclusion de 16 patients pour « intercurrent illness, lack of coopération, or other reasons »	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	25% de données manquantes presque toutes probablement en lien avec la vraie valeur	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?	Même outils de mesure pour tous et même fréquence des visites	PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	« The nature and frequency of side effects did not differ between the 2 active drugs »	PN
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?		NA
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?		NA
Evaluation du risque de biais		Risque faible

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.	Analyse en ITTm et per protocole et uniquement analyse en ITTm présentée	PY
Evaluation du risque de biais		Risque élevé

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Nous considérons que la randomisation et la dissimulation de la séquence d'allocation ont été réalisées dans les règles compte-tenu de l'équipe d'investigation Données démographiques de la population initiale ne suggérant pas de problème de randomisation
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle a priori maintenu Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	11,5% de données manquantes Absence d'information concernant la gestion des données manquantes Données manquantes en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque faible	Méthode de mesure adaptée Double aveugle à priori maintenu Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Risque élevé	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement Deux analyses réalisées : En ITTm et en per protocole. Uniquement l'analyse en ITTm rapportée.
Autres biais		Pas d'informations sur les conflits d'intérêts Pas de méthode d'ajustement du risque alpha rapportée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Carette Année : 1994					
Etude randomisée, en 3 groupes, en schéma parallèle Population à la randomisation : n = 208 patients (amitriptyline 84, cyclobenzaprine 82, placebo 42) Perdus de vue / exclus : 52 (amitriptyline 14, cyclobenzaprine 24, placebo 14) Analyse en ITT					
	Critères de jugement		Mesure du critère	Amitriptyline	Significatif ou non versus Placebo
Critère de jugement principal	50% d'amélioration sur au moins 4 des 6 critères suivants : - douleur, - sommeil, - fatigue, - ressenti global de la fibromyalgie par le patient, - ressenti global de la fibromyalgie par le médecin examinateur, - augmentation de 1kg au score des douleurs musculaires		% de patients traités ayant cette amélioration	1 mois : A : 21% P : 0%	p = 0.002
				3 mois : A : 18% P : 13%	Non significatif, p non donné
				6 mois : A : 36% P : 19%	p = 0.08 non significatif
Critères de jugement secondaires	Douleur		EVA de 0 à 100mm	A : 47 P : 52	Non significatif (p non donné)
	Sommeil		Approximation à partir du graphique	A : 38.5 P : 50	
	Fatigue			A : 50 P : 59.5	
	Amélioration par le patient			A : 48 P : 52.5	
	Amélioration par le médecin			A : 40 P : 48	
	Douleur	Nombre de mots choisis	Mc Gill Pain Questionnaire	1 mois A : 10.6 ± 5.99 // P : 11.7 ± 5.44 3 mois A : 10.1 ± 5.77 // P : 10.6 ± 5.9 6 mois A : 8.87 ± 5.79 // P : 10.2 ± 5.62	Non significatif p > 0.05
Index de notation de la douleur		1 mois A : 21.7 ± 13.7 // P : 26.3 ± 15.0 3 mois A : 21.7 ± 15.0 // P : 22.8 ± 13.5 6 mois A : 19.5 ± 13.5 // P : 21.6 ± 14.4			

	Handicap fonctionnel	Global	Sickness Impact Profile (SIP)	1 mois : A: 15.7 ± 11.3 3 mois : A: 13.9 ± 10.9 6 mois : A: 13.8 ± 11.9	
		Dimension physique		1 mois : A: 8.58 ± 9.00 3 mois : A: 8.80 ± 8.80 6 mois : A: 8.30 ± 8.25	
		Dimension psychologique		1 mois : A: 19.6 ± 16.7 3 mois : A: 15.4 ± 15.0 6 mois : A: 16.2 ± 17.2	
	Index du Handicap fonctionnel	Questionnaire d'évaluation de santé		1 mois : A: 0.60 ± 0.45 3 mois : A: 0.60 ± 0.48 6 mois : A: 0.60 ± 0.49	
	Statut psychologique des patients	Echelle de l'anxiété	Arthritis Impact Measurement Scales	1 mois : 4.60 ± 2.42 3 mois : 4.21 ± 2.10 6 mois : 4.17 ± 2.22	
		Echelle de la dépression		1 mois : A: 2.85 ± 1.71 3 mois : A: 2.55 ± 1.61 6 mois : A: 2.41 ± 1.86	

6. C. Kempenaers 1994: Effect of an antidiacephalon immune serum on pain and sleep in primary fibromyalgia

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles.
Participants	N = 36, femmes âgées de 24 à 56 ans, non ménopausées, non hospitalisées, présentant une fibromyalgie primaire selon les critères de Yunus et Al.
Interventions	Les participants ont reçu au hasard une des 3 molécules ci-dessous : - Un suppositoire d'Antidiacephalon immune serum " SER282 " (20 mg/ml) trois fois par semaine et un comprimé de placebo le soir. - Un suppositoire de placebo trois fois par semaine et un comprimé de 50 mg d'Amitriptyline le soir. - Un suppositoire de placebo trois fois par semaine et un comprimé de placebo le soir
Outcomes	Pas de distinction entre le critère de jugement principal et les critères secondaires : “Évaluer l'effet du SER282 par rapport à l'amitriptyline et au placebo sur les symptômes et la physiologie du sommeil” - La qualité globale du sommeil évaluée sur une EVA de 0 à 100 - Le repos matinal et la fatigue diurne notés de 0 à 3 - Une Polysomnographie a été réalisée 4 nuits en phase initiale et 1 nuit en phase finale, incluant un EEG, un EMG, une thermistance naso buccale, un enregistrement thoraco-abdominal par électrodes pour enregistrer les apnées/hypopnées et un ECG. - Le nombre de points sensibles "Tender points number" - La douleur localisée : score total des douleurs au niveau cervical, scapulaire, lombaire, pelvien et aux membres supérieurs et inférieurs. - Les polyalgies : score de 0 à 3 - Les symptômes associés : score total pour les troubles du sommeil, asthénie matinale, fatigue, crampes, dysménorrhée, la constipation, les céphalées, les vertiges, le gonflement des doigts. - Les troubles du sommeil - L'auto-évaluation globale de la douleur sur une EVA de 0 à 100mm - Echelle d'Hamilton pour la dépression - La tolérance et l'efficacité ont également été notées respectivement de 0 à 2 et de de -1 à + 3
Notes	Critères d'exclusion : la présence d'un syndrome inflammatoire, un dosage de facteur rhumatoïde positif ou un diagnostic de polyarthrite rhumatoïde, un trouble psychiatrique sévère ou toute pathologie métabolique ou endocrinienne. Étude monocentrique, durée 8 semaines Analyse en per protocole Participants analysés : N =23 SER282 n = 9, Amitriptyline n = 6, placebo n = 8 Perdus de vue/exclusion : N = 13

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« according to randomization and double-blind design »	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	Absence d'informations supplémentaires	NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	"Due to randomization bias, these patients showed more associated symptoms, a more frequent history of affective disorders and use of antidepressants [...]"	PY
Evaluation du risque de biais		Risque élevé

Domaine 2 : Biais dû aux écarts par rapport aux interventions envisagées (effet de l'adhésion à l'intervention,)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Absence d'informations concernant les effets indésirables spécifiques de l'amitriptyline.	NI
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?	Remarque : 2 fois plus d'abandons dans le groupe amitriptyline.	NI
2.3 Si Y / PY / NI à 2.1 ou 2.2 Les interventions non protocolaires importantes étaient-elles équilibrées entre les groupes d'intervention ?		Non concerné
2.4 [Le cas échéant] Y a-t-il eu des échecs dans la mise en œuvre de l'intervention qui auraient pu affecter les résultats ?	Nombre de patients n'ayant pu bénéficier d'EEG par groupe inégal Pas d'explication quant aux raisons de la non-réalisation des EEG sur tous les participants ayant fini l'étude avec un contexte de déséquilibre de facteur pronostic dans ce groupe péjoratif	NI

	(plus de patients atteints de dépression majeure dans le groupe SER282 que dans les 2 autres groupes)	
2.5 [Le cas échéant] Y a-t-il eu un non-respect du schéma d'intervention assigné qui aurait pu affecter le résultat des participants ?	2 fois plus d'abandon dans le groupe amitriptyline (50% = 6 patients) que SER282 (25% = 3 patients), 1,5 fois plus que placebo (23% = 4 patients)	PY
2.6 Si N / PN / NI à 2.3 ou Y / PY / NI à 2.4 ou 2.5 : Une analyse appropriée a-t-elle été utilisée pour estimer l'effet du respect de l'intervention ?	Absence d'analyse rapportée	PN
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes

Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	13 perdus de vue, 0 exclus sur 36 patients => 36.1% données manquantes Et en plus seul 19 patients ont eu les EEG sur 36 => 17/36 données manquantes : 47.2% données manquantes sur EEG	PY
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas d'informations concernant la gestion des données manquantes	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	« treatment inefficacy or aggravation symptoms (n=9), side effects (n= 4), or for unknown reason (n= 3) » 10 données manquantes en lien sûr avec vraie valeur => 27.7% de données manquantes en lien sûr avec vraie valeur et 3 peut-être, pas d'info.	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	2 fois plus d'abandon groupe amitriptyline (50% = 6 patients) que SER282 (25% = 3 patients), 1,5 fois plus que placebo (23% = 4 patients) et beaucoup trop de données manquantes en lien avec vraie valeur.	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?	Non	PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?	Non	PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Doute sur le maintien du double aveugle	NI
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Evaluateur : patient et/ou médecin Critères de jugements subjectifs	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produit ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	- Pas le résultat de la tolérabilité - Pas le résultat des échelles anxiété et dépression, juste le p pour anxiété - Juste le p pour sleep disturbance mais pas de résultat	PY
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Risque élevé

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque élevé	Peu d'informations concernant la séquence de randomisation ainsi que sur la dissimulation de la séquence d'allocation Les données démographiques de la population initiale suggèrent un problème de randomisation
Biais dû aux déviations par rapport à l'intervention prévue (adhésion à l'intervention)	Risque élevé	Absence d'informations pour statuer sur le double aveugle (Effets indésirables spécifiques de l'amitriptyline) Absence d'information concernant la survenue d'échecs de mise en œuvre des interventions Manque d'adhésion à l'intervention Pas d'analyse appropriée pour estimer l'effet de l'adhésion
Biais dû aux données manquantes	Risque élevé	47,2% de données manquantes sur les résultats Absence d'information concernant la gestion des données manquantes Données manquantes en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Maintien du double aveugle non statué Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Risque élevé	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement Résultats mesurés non intégralement rapportés
Autres biais		Pas d'informations sur les conflits d'intérêts Pas de méthode d'ajustement du risque alpha mentionnée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Kampenaers Année : 1994					
Etude randomisée parallèle 3 bras : placebo, amitriptyline, SER282 Analyse en per protocole 36 patients randomisés, 13 perdus de vue / exclu, Examen clinique final : SER282 n = 9, Amitriptyline (A) n = 6, placebo (P) n = 8 Examen EEG final : SER 282 n = 6, Amitriptyline (A) n = 6, placebo (P) n = 7 Pas de comparaison de l'amitriptyline contre placebo					
	Critère de jugement	Valeur rapportée	Amitriptyline	Placebo	Significativité P (t=time effect) P (g = group effect) P tg (time x groupe)
Critères de jugements sur l'examen clinique final (pas de distinction primaire / secondaire)	Nombre de points sensibles (0 -> 18)	Moyenne (déviati on standard)	T0 = 16 4°S = 11 (4) 8°S = 10 (5)	T0 = 16 (2) 4°S = 14 (4) 8°S = 11 (5)	Pt = 0.0001 Pg = ns Ptg = 0.04
	Sévérité de la douleur localisée (0 ->18)		T0 = 12 (1) 4°S = 7 (5) 8°S = 9 (4)	T0 = 10 (2) 4°S = 8 (3) 8°S = 6 (3)	Pt = 0.004 Pg = ns Ptg = 0.035
	Sévérité des polyalgies (0 -> 3)		T0 = 2.5 (0.6) 4°S = 1.0 (1.3) 8°S = 1.7 (1.0)	T0 = 1.9 (0.6) 4°S = 1.5 (0.9) 8°S = 1.1 (1.0)	Pt = 0.001 Pg = ns Ptg = 0.045
	Douleur globale (0 -> 100)		T0 = 57 (23) 4°S = 42 (27) 8°S = 32 (31)	T0 = 65 (20) 4°S = 50 (23) 8°S = 37 (28)	Pt = 0.012 Pg = ns Ptg = ns
	Symptômes associés (0 -> 27)		T0 = 13 (3) 4°S = 7 (4) 8°S = 5 (3)	T0 = 14 (3) 4°S = 10 (4)	Pt = 0.001 Pg = 0.025 Ptg = ns

				8°S = 4 (4)	
	Qualité globale de sommeil (VAS 0- >100)		T0 = 30 (8) 8°S = 47 (30)	T0 = 58 (7) 8°S = 60 (22)	Pt = Pg = Ptg = ns
	Fatigue diurne (0 - > 3)		T0 = 1.5 (0.6) 8°S = 1.2 (0.5)	T0 = 2.0 (1.0) 8°S = 2.7 (1.5)	Pt = Pg = ns Ptg = 0.05
Critères de jugement sur EEG (pas de distinction primaire / secondaire)	Index d'éveil, n/h		T0 = 7 (2) 8°S = 10 (3)	T0 = 8 (3) 8°S = 6 (2)	Ptg = 0.029
	Stage 4		T0 = 28 (12) 8°S = 11 (14)	T0 = 41 (29) 8°S = 37 (21)	Ptg = 0.55
	Stage 2		T0 = 240 (33) 8°S = 237 (36)	T0 = 208 (40) 8°S = 245 (63)	Pg = 0.059
	Sommeil sur la latence définie (minutes)		T0 = 16 (24) 8°S = 22 (20)	T0 = 15 (11) 8°S = 12 (5)	Pt = Pg = Ptg = ns
	Index d'efficacité (temps de sommeil / temps au lit)		T0 = 89.2 (4.5) 8°S = 82.5(6.5)	T0 = 85.1 (6.5) 8°S =87.1 (7.1)	Pt = Pg = Ptg = ns
	Stage shifts index (n/h)		T0 = 44 (10) 8°S = 49 (10)	T0 = 39 (13) 8°S =36 (12)	Pt = Pg = Ptg = ns
	Stage 1, min		T0 = 46 (10) 8°S = 56 (9)	T0 = 40 (16) 8°S = 44 (19)	Pt = Pg = Ptg = ns
	Stage 3, min		T0 = 42 (9) 8°S = 28 (12)	T0 = 32 (8) 8°S = 33 (12)	Pt = Pg = Ptg = ns
	Sommeil REM		T0 = 88 (17) 8°S = 89 (29)	T0 = 89 (42) 8°S = 90 (33)	Pt = Pg = Ptg = ns

	Latence REM (min)		T0 = 69 (19) 8°S = 78 (31)	T0 = 82 (40) 8°S = 79 (31)	Pt = Pg = Ptg = ns
	Durée totale de sommeil (min)		T0 = 443 (42) 8°S = 421 (65)	T0 = 409 (74) 8°S = 449 (64)	Pt = Pg = Ptg = ns
Résultats rapportés dans le texte uniquement et incomplets	<u>Sleep disturbance (0->3)</u> : « a differential evolution of symptoms with treatment was suggested for TP (p = 0.04), total localized pain (p = 0.035), polyalgia (p = 0.045) and sleep disturbance (p = 0.023) » mais pas de résultat donné pour sleep disturbance, que le « p » <u>Clinician-ratings of mood</u> : hamilton scale of depression and anxiety : « also anxiety decreased significantly (p = 0.001) ; depression score – which was modest at baseline – remained stable. » mais pas de résultat donné ni pour anxiété ni dépression, et pas de « p » pour dépression <u>Morning restfulness</u> : résultats non rapportés Tolerance : résultats non rapportés <u>Efficacy</u>: « concerning the global assessment of treatment response, patients reported a significant positive effect of SER282 and AMI from week 4 on, with PL providing benefit at week 8 only (all p values < 0.05). Clinician-rated improvement was only significant for AMI from week 4 on (p< 0.005); PL was effective at week 8 (p < 0.05), while no change was noted with SER282 at week 4 or 8.				

7. S. Carette 1995: Sleep electroencephalography and the clinical response to amitriptyline in patients with fibromyalgia

a) Caractéristiques de l'étude :

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en crossover avec 2 bras d'intervention.
Participants	N = 22 Adultes de plus de 18 ans, remplissant les critères de l'American College of Rheumatology de 1990 pour la classification de la fibromyalgie. De plus, les patients devaient avoir un score supérieur ou égal à 4 sur au moins une des deux EVA auto-administrées de 10 cm (Une évaluant la douleur et une évaluant le bilan global des symptômes de la fibromyalgie). Les participants devaient avoir des résultats normaux concernant le dosage de la TSH et des CPK, ainsi qu'un test de vitesse de sédimentation érythrocytaire normal.
Interventions	Les participants ont reçu au hasard : - Soit Amitriptyline 25mg au coucher - Soit un placebo au coucher
Outcomes	<u>Critère de jugement principal :</u> Etude de l'impact d'un type d'anomalies du sommeil (alpha non rapide mouvement oculaire) sur les symptômes de la fibromyalgie. <u>Critères de jugement principal :</u> - La douleur sur une EVA de 0 à 10mm - La fatigue sur une EVA de 0 à 10mm - Le sommeil sur une EVA de 0 à 10mm - L'évaluation globale de la fibromyalgie par le patient - L'évaluation globale de la fibromyalgie par le médecin - Le score myalgique total "total myalgic score" par la somme des mesures des points sensibles de la fibromyalgie - Amélioration significative d'au moins 4 des 6 critères chez les patients : - 50 % d'amélioration de la douleur, du sommeil, de la fatigue, de l'évaluation globale du patient ou de l'évaluation globale du médecin
Notes	Critères d'exclusion : Les patients étaient exclus s'ils présentaient des signes de maladie rhumatismale inflammatoire, endocrinienne non traitée, un trouble neurologique, infectieux, musculaire ou osseux. Les patients ayant des antécédents de glaucome, de rétention urinaire, d'infarctus du myocarde, d'arythmie, d'insuffisance cardiaque congestive ou d'apnée du sommeil. Les anti-inflammatoires non stéroïdiens, hypnotiques devaient avoir été arrêtés 2 semaines avant l'entrée dans l'essai, tandis que les antidépresseurs devaient avoir été arrêtés 1 mois auparavant.

	Le seul traitement hors étude autorisé pendant l'essai était le paracétamol. Une lettre a été envoyée au médecin traitant expliquant l'importance de limiter les interventions concomitantes. A noter que 4 des patients avaient déjà participé à l'étude de Carette 1994. Etude monocentrique Durée totale : 16 semaines Période traitement : 8 semaines Période de wash-out : absence Analyse en ITTm Participants analysés N = 20 Perdus de vue / exclus N = 2 après la première période Résultat significatif pour l'amitriptyline versus placebo pour le critère de jugement principal (p=0,02). Résultat significatif pour l'amitriptyline versus placebo avec amélioration de : - La douleur (p<0.05) - La fatigue (p<0.05) - Le sommeil (p<0.05) - L'évaluation globale de la fibromyalgie par le patient et le médecin (p<0.05)
--	---

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« The order of treatment was generated using a table of random numbers »	PY
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	Equipe expérimentée (cf Carette 1986 et 1996) Etude monocentrique mais réalisée dans un Centre hospitalo universitaire.	PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	Présence d'un tableau de caractéristiques de base sur la population de base globale.	NI
Evaluation du risque de biais		Risque faible

Biais dû à l'effet « carry-over »		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	Le ratio utilisé semble être 1 : 1 car « each patient was randomly assigned to receive either amitriptyline 25 mg/j, 1 hour before bedtime or an identical-appearing inert placebo »	PY
S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?		Non concerné
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	« No wash-out treatment period was instituted between the two 8- week treatment periods since it was not expected that amitriptyline would affect patients' symptoms and sleep parameters evaluated 2 months after cessation of this treatment »	PY
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Absence d'informations concernant les effets secondaires spécifiques de l'Amitriptyline.	NI
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		NI
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné

2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITTm : 2 perdus de vue sur 22 => 9.1% de données manquantes	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires-	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	2 pdv sur 22 => 9.1% de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Absence d'information concernant la gestion des données manquantes	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Pas d'informations sur les raisons d'abandon de l'étude donc possible	NI
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	Idem 3.3	NI
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?	Même méthode de mesure dans chaque groupe	PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Evaluateurs : patient et/ou médecin Doute concernant le maintien de l'aveugle. Effets secondaires spécifiques de l'amitriptyline non rapportés.	NI
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critères de jugements subjectifs Evaluateur : patient et/ou médecin	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produit ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données ?		PN
5.4 : : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Nous considérons que la randomisation et la dissimulation de la séquence d'allocation ont été réalisées dans les règles compte tenu de l'équipe d'investigation Données concernant la population de base globale et non par groupe
Biais dû à l'effet carry-over	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Pas assez d'informations pour statuer concernant le maintien du double aveugle Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	9.1 % de données manquantes Absence d'information concernant la gestion des données manquantes Absence d'informations concernant les raisons des données manquantes
Biais dû à la mesure de l'outcome	Risque élevé	Maintien du double aveugle non statué Evaluateur : patient et/ou médecin Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai. Protocole a priori non enregistré préalablement Absence de sélection de résultats ou d'analyse Les résultats pour les deux périodes sont disponibles
Autres biais		Absence d'informations concernant les conflits d'intérêts Méthode d'ajustement du risque alpha par la méthode de Tukey
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Carette Année: 1995					
Etude randomisée en crossover 2 bras : Amitriptyline et Placebo Durée : 2 périodes de 8 semaines Absence de wash-out Analyse en ITTm Randomisation : 22 patients, Fin de traitement : 20 patients Perdus de vue : N = 2					
	Critère de jugement	Mesure du critère	Amitriptyline	Placebo	Significatif ou non vs placebo
Critères de jugement secondaires	Amélioration significative d'au moins 4 des 6 critères : 50 % d'amélioration - douleur - sommeil - fatigue - évaluation globale -> par le patient -> par le médecin	% patients répondeurs	27.3%	0	P = 0.02
Critères de jugement secondaires	Douleur (patient)	VAS (0 -> 10cm) Moyennes ± écart type	5.07 ± 3.22	7.13 ± 2.41	P < 0.05
	Fatigue (patient)		5.62 ± 3.07	7.64 ± 1.80	P < 0.05
	Sommeil (patient)		3.93 ± 3.14	6.51 ± 2.69	P < 0.05
	Evaluation globale (patient)		5.47 ± 3.03	7.11 ± 2.14	P < 0.05
	Evaluation globale (médecin)	4.18 ± 2.81	6.36 ± 1.59	P < 0.05	
	Total Myalgic Score (médecin) (Somme des points sensibles 0 -> 10)	Moyenne ± écart type	3.45 ± 1.16	3.22 ± 0.86	
Résultats des analyses des patients en fonction de « alpha rating sur EEG »					
Critère de jugements principaux		Alpha normale (n = 14)	Alpha anormale (n = 8)	Différence (95%CI)	
	% de répondeurs à l'amitriptyline	28.6	25.0		
	Douleur	7.16 ± 1.49	7.06 ± 2.58	0.10 (-1.7, 1.90)	
	Fatigue	7.73 ± 1.85	8.04 ± 1.80	- 0.31 (-2.00, 1.38)	
	Sommeil	6.76 ± 3.2	8.76 ± 1.00	- 2.00 (-4.45, 0.45)	
	Evaluation globale (patient)	7.09 ± 1.68	7.58 ± 1.78	- 0.49 (-2.08, 1.1)	
	Evaluation globale (médecin)	6.26 ± 1.41	6.67 ± 2.16	- 0.41 (-1.99, 1.17)	
	Score myalgique total	3.21 ± 0.74	3.12 ± 0.79	0.09 (- 0.61, 0.79)	

8. Don. L. Goldenberg 1996: A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia

a) Caractéristiques de l'étude:

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en cross-over avec 4 bras d'intervention.
Participants	N = 31 Adultes âgés de 18 à 60 ans remplissant les critères de l'American College of Rheumatology de 1990. Par ailleurs, chaque patient devait avoir un score supérieur ou égale à 30 (maximum 100) sur l'échelle visuelle analogique (EVA) de la douleur lors de la visite d'accueil.
Interventions	Les participants ont reçu soit : - De l'Amitriptyline au coucher à la dose de 25 mg + un Placebo le matin Soit - De la Fluoxétine le matin à la dose de 20 mg + un Placebo au coucher Soit - De l'Amitriptyline au coucher à la dose de 25 mg + de la Fluoxétine le matin à la dose de 20 mg Soit - Un Placebo le matin + un Placebo au coucher
Outcomes	Le critère de jugement principal semble être l'amélioration globale par le patient car elle a servi au calcul du nombre de sujets nécessaires pour l'étude. <u>Critères de jugements secondaires évalués par le médecin :</u> - Un "Tender points score " obtenu en additionnant les scores à la palpation de chaque point sensible - Le bien être global du patient sur une EVA de 0 à 100mm - Les effets indésirables <u>Critères de jugements secondaires par le patient :</u> - La Douleur sur une EVA de 0 à 100 mm - Les troubles du sommeil sur une EVA de 0 à 100 mm - La fatigue sur une EVA de 0 à 100 mm - La sensation d'être reposé au réveil sur une EVA de 0 à 100 mm - Le bien être global sur une EVA de 0 à 100 mm - Le Fibromyalgia Impact Questionary (FIQ) - Le Beck Depression Inventory (BDI)
Notes	Critères d'exclusion : antécédents de maladie systémique, y compris cardiaque, rénale, maladie hématologique ou hépatique ; le refus d'interrompre tous les médicaments agissant sur le

système nerveux central, les anti-inflammatoires non stéroïdiens et les analgésiques (autres que le doliprane) pendant au moins

1 semaine avant l'étude.

Afin d'exclure les patients souffrant de dépression majeure, chaque patient devait avoir un score de 518 sur l'échelle d'évaluation d'Hamilton pour la dépression.

Etude monocentrique

Durée de l'étude : 30 semaines (4 périodes de 6 semaines séparées par 2 semaines de wash-out)

Période de traitement : 6 semaines

Période de wash-out : 2 semaines

Analyse en ITTm

Participants analysés N = 19

Perdus de vue / exclus N = 12

L'amitriptyline n'est pas comparée au placebo, résultats donnés en intragroupe significatifs pour l'amitriptyline pour :

- Le bien être global mesuré par le patient et le médecin
- La douleur
- Les troubles du sommeil
- Le FIQ

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« using randomization tables, the pharmacy assigned patients to 1 of the following 4 regimens » « the order of treatment was generated from a table of random numbers. Randomization was performed in the hospital pharmacy » Donc randomisation centralisée correcte	Y
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		Y
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« There were no statistically significant differences in baseline characteristics and baseline responses between the 19 patients who completed the study and the 12 who dropped out (table 1), except that those who dropped out were more likely to be working. »	NI
Evaluation du risque de biais		Risque faible

Risque de biais dû à l'effet carry-over		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
S.1 Est-ce que le nombre de participants dans chaque période dans chaque bras était équivalent ?	Ratio de randomisation : 1 :1 :1 :1	PY
S2. Si N, PN, NI à question S.1 : Est-ce que la période d'effet est prise en compte dans l'analyse ?		Non concerné
S3. Est-ce qu'il y a eu assez de temps pour que les effets « carry-over » disparaissent entre les 2 interventions ?	« We used the score at the end of each 6-week treatment period as outcome in repeated analysis of variance[...] Because of the 8-week interval between outcome measures, including the 2-week wash-out periods, carry-over effects were not expected »	PY
Evaluation du risque		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur était assignée pendant l'essai ?	« All tablets were identical in appearance » Schéma complexe à 4 bras en cross over Allocation randomisée centralisée « there were no significant differences in the type of adverse reaction in patients during of the 4 trial periods »	PN
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PN
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		Non concerné
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à	Analyse en ITTm “ All data on patients who completed at least 1 treatment period were analyzed for outcome, but patients who failed to complete a period had no outcome that could be measured. “ 12 perdus de vue soit 38.7% de données manquantes	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	Analyse en ITTm avec plus de 5% de perdus de vue	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	31 patients randomisés, 12 perdus de vue =>12/31 = 38.7% de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Absence d'information concernant la gestion des données manquantes.	PN
3.3 <u>Si N / PN à 3.2</u> Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	“5 patients cited an adverse reaction to the medication as reason for dropping out of the trial 4 patients [...] cited increased FM symptoms 3 [...] withdrew for reasons other than FM symptoms or adverse side effects to the medication.”	PY
3.4 <u>Si Y/ PY / NI à 3.3</u> : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	Critères d'abandon de l'étude en lien avec vraie valeur : effets indésirables et majoration des symptômes	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Evaluateur = médecin et patients Double aveugle maintenu	PN
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?		NA
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?		NA
Evaluation du risque de biais		Risque faible

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
5.4 : : Le résultat présenté, initialement prévu sur deux périodes, est-il non disponible du fait d'un effet carry-over ?		N
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Séquence d'allocation aléatoire et dissimulation de la séquence réalisée Données concernant la population de base globale et non par groupe
Biais dû à l'effet carry-over	Risque faible	Nombre de participants dans chaque bras dans chaque période équivalent Absence d'effet carry-over identifié
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	A priori maintien du double aveugle Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	38.7 % de données manquantes Absence d'information concernant la gestion des données manquantes Causes des données manquantes en lien avec sa vraie valeur
Biais dû à la mesure de l'outcome	Risque faible	Maintien du double aveugle Evaluateur : patient et/ou médecin Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai. Protocole a priori non enregistré préalablement Absence de sélection de résultat ou d'analyse Les résultats pour les deux périodes sont disponibles
Autres biais		Absence d'informations concernant les conflits d'intérêts Méthode d'ajustement du risque alpha rapporté dans l'essai 1986, a priori également réalisé
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Don Goldenberg				Année : 1996			
Etude randomisée en crossover 4 bras d'intervention Amytriptyline + Placebo, Amitriptyline + Fluoxétine, Fluoxétine + Placebo, Placebo + Placebo Analyse en ITTm Participants N = 31 Période de traitement : 6 semaines Période de wash-out : 2 semaines A priori amitriptyline non comparée au Placebo dans les tableaux de l'étude							
	Critère de jugement		Mesure du critère	Valeur rapportée	Amitriptyline	Placebo	Significatif ou non en intragroupe amitriptyline
Critère de jugement principal	Patient	Amélioration globale	EVA 0-->100		61.6 (29.5)	76.8 (24.8)	P = 0.02
Critères de jugement Secondaires	Médecin	Score de points sensibles		Moyenne (écart type)	18.0 (7.2)	19.0 (7.5)	P = 0.3
		Bien être global	EVA 0->100 mm		64.2 (25.2)	74.7 (19.9)	P = 0.04
	Patients	Douleur	EVA 0->100 mm		64.4 (28.3)	81.5 (16.5)	P = 0.02
		Troubles du sommeil			57.0 (34.8)	74.6 (23.9)	P < 0.001
		Fatigue			67.7 (29.9)	73.7 (25.1)	P = 0.26
		Sensation d'être reposé au réveil			67.7 (29.9)	73.7 (25.1)	P = 0.26
		Fibromyalgia Impact Questionary (FIQ)			52.3 (22.9)	58.5 (17.1)	P = 0.03
		Beck Depression Inventory (BDI)			8.7 (6.0)	9.3 (6.5)	P = 0.52

9. F. Ginsberg 1996 : A randomized placebo-controlled trial of sustained-release amitriptyline in primary fibromyalgia

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles.
Participants	N = 51 Adultes de plus de 18 ans, remplissant les critères de l'American College of Rheumatology de 1990 pour la classification de la fibromyalgie. Les participants devaient présenter une fibromyalgie primaire en dehors de toute affection rhumatologique et avoir un antécédent de douleur généralisée depuis au moins 3 mois et une douleur dans au moins 11 des 18 points sensibles spécifiques à la palpation digitale.
Interventions	Les participants ont reçu soit : - De l'Amitriptyline en libération prolongée à la dose de 25 mg par jour. Soit - Un Placebo
Outcomes	Critère primaire d'efficacité, définie par au moins 4 critères : Au moins 50 % d'amélioration dans l'évaluation globale du patient. Au moins 50 % d'amélioration dans l'évaluation globale de l'examineur. Au moins 50 % d'amélioration de la douleur. Au moins 25% de réduction du score de points sensibles de la fibromyalgie. Les critères secondaires d'efficacité incluent les analyses entre groupes et en inter groupe des mesures d'efficacité individuels : - Evaluation globale des symptômes par le patient et par l'investigateur sur une EVA de 0 à 10mm - Evaluation de la douleur par le patient sur une EVA de 0 à 10 mm - Nombre de points sensibles positifs du fibromyalgie noté de 0 à 18 d'après Wolfe et al, évalué par un dolorimètre de 4 kg. - Score de points sensibles de la fibromyalgie noté de 0 à 72 évalué par un dolorimètre de 4 kg. - Evaluation des troubles du sommeil sur la base de 3 questions avec des réponses notées sur une EVA de 0 à 10 mm : A) Avez-vous des difficultés à dormir ? B) Vous réveillez vous fatigué ou non reposé ? C) Comment vous sentez vous le matin au réveil ? - Evaluation de la fatigue en posant la question :

	Si vous êtes fatigué, l'êtes-vous rarement, souvent ou de manière habituelle? Et en évaluant l'importance de la fatigue sur une EVA de 0 à 10mm. - Evaluation de la durée de la raideur matinale Les critères secondaires d'efficacité incluent les analyses entre groupes et en inter groupe des mesures d'efficacité individuels. Les effets secondaires ont été recueillis sur des plaintes spontanées des patients ou sur l'observation d'effets secondaires. Un examen sanguin biologique et biochimique a été réalisé initialement pour identifier une pathologie intercurrente.
Notes	Critères d'exclusion : Les participants étaient exclus s'il y avait une grossesse potentielle, une grossesse confirmée, un allaitement en cours, un glaucome à angle fermé, un risque de rétention urinaire, un antécédent de syndrome coronarien, d'épilepsie ou d'insuffisance cardiaque, d'arythmie cardiaque ou de troubles de la conduction cardiaque. Les participants étaient également exclus s'ils n'arrivaient pas à interrompre à l'entrée de l'étude : un traitement contenant de la vitamine D, du magnésium, un traitement par myorelaxants, antalgiques (hormis le paracétamol), les antis inflammatoires, les antidépresseurs, les somnifères, tranquillisants ou tout autre médicament ayant une action sur le système nerveux central. Les personnes ayant utilisé de l'amitriptyline durant les 6 mois précédant l'essai étaient également exclues, les personnes ayant utilisé des inhibiteurs de la monoamine oxydase durant les 3 semaines précédant l'essai, ainsi que les personnes présentant une hypersensibilité à l'amitriptyline, aux antidépresseurs tricycliques ou au paracétamol. Etude monocentrique Durée de l'étude : 8 semaines Analyse en ITTm Finalement 46 participants analysés Amitriptyline N = 24, Placebo N = 22 Perdus de vue/exclus N = 5 : 2 amitriptyline et 3 placebo Résultat significatif pour l'amitriptyline pour le critère de jugement principal à 4 et 8 semaines. A 8 semaines : résultats significatifs concernant la fatigue majorée pour le groupe Placebo Amélioration significative de la fatigue au réveil pour le groupe Amitriptyline versus Placebo à semaine 4 comparée au début, et de la fatigue à semaine 8 comparée au début.

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« At the entry into the study each patient was randomly assigned to receive either sustained-release amitriptyline [...] or placebo witch were identical to amitriptyline capsules, except that placebo capsules did not contain the active princip. Active and placebo capsules were provided by sanofi Pharma, Brussels, Belgium » Nous en déduisons que le laboratoire a uniquement fourni les médicaments et n'a procédé à une randomisation centralisée.	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		NI
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« there were no significant differences between the treatment groups with respect to the patients' demographic and clinical characteristics at study entry [tables 1 and 2] » Caractéristiques de la population finale présentées.	NI
Evaluation du risque de biais		Quelques inquiétudes

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	“Adverse effects were reported in 7 patients (29%) of group A and in none of group P (p<0,01) “ 3 bouches sèches dans le groupe amitriptyline ce qui correspond à une effet indésirable spécifique.	PY
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PY
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN

2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse ITTm avec 5 pdv exclus d'analyse soit 5/51 = 9.8% de données manquante	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	Analyse non appropriée avec plus de 5% de perdus de vue	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	17.6 % de perte de données à la randomisation : 5 perdus de vue sur 51 patients après la phase initiale “5 patients of whom 2 were randomized to A and 3 to P where lost to follow up after the inclusion visit. Since non efficacy or safety data was available for these patients, the analysis based on the ITT principe was confined to 46 patients.” Par la suite 1 abandon dans le groupe amitriptyline et 3 dans le groupe placebo.	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas de méthode d'estimation des données manquantes pour les 5 premiers perdus de vue. Méthode d'estimation pour les 3 perdus de vue durant l'essai mais non rapportée.	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Pas d'informations sur les raisons d'exclusion donc on ne peut exclure un possible lien avec leur état de santé	NI
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?		NI
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Double aveugle non maintenu	PY
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critère de jugement subjectif Evaluateur : patient et/ou médecin	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez s'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Quelques inquiétudes	Peu d'informations concernant la randomisation Caractéristiques de la population initiale non fournies
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Double aveugle a priori non maintenu Analyse en ITTm avec plus de 5% de données manquantes
Biais dû aux données manquantes	Risque élevé	9.8 % de données manquantes Absence d'information concernant la gestion des données manquantes Absence d'informations sur les causes de données manquantes
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Double aveugle à priori non maintenu Critères de jugements subjectifs Evaluateur : patient et/ou médecin
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : Protocole non enregistré préalablement Absence de sélection de résultat ou d'analyse
Autres biais		Pas d'informations sur les conflits d'intérêts Pas de méthode d'ajustement du risque alpha rapportée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Ginsberg Année : 1996					
Essai randomisé en parallèle 2 bras Amitriptyline versus Placebo Analyse en ITTm Participants N = 51 Fin de traitement : 46 (24 amitriptyline, 22 placebo)					
	Critère de jugement	Mesure du critère	Valeur rapportée	Amitriptyline	Résultat Significatif ou non versus Placebo
Critère de jugement primaire	Réalisation de 3 critères sur 4 : -Au moins 50 % d'amélioration dans l'évaluation globale par patient -Au moins 50 % d'amélioration dans l'évaluation globale par médecin -Au moins 50 % d'amélioration de la douleur -Au moins 25% de réduction du score de points sensibles		Pourcentage de répondeur dans chaque groupe %	4°S A = 5 patients répondeurs [21%, 95% CI 7.1 – 42.2%] P= 0 patient 8°S A =14 patients répondeurs [58%, [36.6 – 77.9%] P = 0 patient	4°S P = 0.05 8°S P < 0.001
Critère de jugement secondaire	Evaluation Globale de la fibromyalgie EVA (0 -> 10cm)	Par le patient	Moyenne et écart type	T0 A : 7.7 ±1.3 P : 7.0 ± 1.6 4°S : A: 5.2 ± 2.1 P : 7.3 ± 1.2 8°S :	Non comparé au placebo mais interaction traitement

				A: 3.9 ± 2.3 P : 6.8 ± 1.8	et temps en intra groupe P < 0.001 pour le groupe amitriptyli ne
		Par le médecin		T0 : A : 7.0 ± 1.3 P : 6.2 ± 1.5 4°S : A : 4.4 ± 2.2 P : 5.9 ± 1.3 8°S : A: 3.5 ± 2.1 P : 5.9 ± 1.7	
	Douleur EVA (0 -> 10cm)			T0 : A : 7.3 ± 1.4 P : 7.1 ± 1.4 4°S : A: 5.0 ± 2.1 P: 7.1 ± 1.3 8°S : A: 3.8 ± 2.4 P : 7.0 ± 1.3	
	Points sensibles	Nombre de points sensibles (0 ->18)		T0 : A : 14.8 ± 2.1 4°S : A: 11.8 ± 3.8 8°S : A: 10.2 ± 4.8	
		Score (0 ->72)		T0 : A: 23.6 ± 7.5 4°S : A: 16.8 ± 9.4 8°S : A: 13.6 ± 9.6	
	Troubles du sommeil	Sensation au réveil (EVA 0- >10cm)		T0 : A: 6.3 ± 2.0 P: 5.9 ± 2.2 4°S : A: 4.1 ± 2.5 P: 5.4 ± 2.5 8°S : A: 3.2 ± 2.6 P: 5.3 ± 2.3	
			Amélioration des patients (Pourcentage)	4°S 50% groupe A vs 9% groupe P 8°S	4°S P= 0.008 Significatif vs placebo 8°S

				75% groupe A vs 9% groupe P		P< 0.001 Significatif versus placebo
		Fatigue au réveil (Jamais / rarement / souvent / toujours)		Jama is / rare ment	Souven t/ toujour s	Non significatif versus placebo
				T0 : A= 5 4°S : A= 11 8°S : A= 14	T0 : A= 19 4°S : A=13 8°S : A=10	
		Troubles du sommeil (0 -> 10cm)		T0 : A: 5.2 ± 2.5 P :5.4 ± 3.0 4°S : A: 3.3 ± 2.8 P : 5.3 ± 2.9 8°S : A:2.6 ± 3.1 P: 5.1 ± 3.0		Non comparé au placebo mais interaction traitement et temps en intra groupe P = 0.003 pour le groupe amitriptyli ne
	Fatigue	Rarement / souvent / toujours		Jamais / rarement	Sou ven t/ touj our s	P = 0.040 Significatif contre placebo
				T0 : A=0 P=5 4°S : A= 8 P= 6 8°S : A=11 P=7	T0 : A= 24 P= 17 4°S : A= 16	

					P=1 6 8°S : A= 13 P=1 5	
			Amélioration des patients (Pourcentage)	5°S 46% groupe A vs 18% groupe P 8 ° S : 58% patients améliorés groupe A vs 18% groupe P	Comparé au Placebo 5° S p= 0.094 8°S p = 0.014	
		EVA (0 -> 10cm)		T0 : A: 7.3 ± 1.5 P : 6.7 ± 2.0 4°S : A: 5.1 ± 1.9 P : 6.2 ± 2.4 8°S : A: 3.8 ± 2.5 P: 5.9 ± 2.2	Non comparé au placebo mais interaction traitement et temps en intra groupe P = 0.001 pour le groupe amitriptyline	
	Durée de la raideur matinale	Minutes		T0 : A: 32.5 ± 29.5 P : 28.0 ± 27.9 4°S : A: 16.7 ± 22.1 P : 25.9 ± 25.0 8°S : A: 9.7 ± 17.7 P: 22.5 ± 22.1	Non comparé au placebo mais interaction traitement et temps en intra groupe P = 0.006 pour le groupe amitriptyline	

Autres	Compliance au traitement	Compte des comprimés retournés	« No relevant differences between the groups in the overall good compliance with study treatment »
	Prise de paracétamol		« In each treatment group, 7 patients took paracetamol as rescue medication and the number of tablets of paracetamol taken was similar in both groups
	Prise de médicaments interdits		« 3 patients in each group took prohibited medication benzodiazepine [...] antiinflammatory agents »
	Tolérance clinique	Plaintes spontanées des patients	« Adverses events were reported in seven patients [29%] of group A and none in group P [fisher's exact test, $p = 0.010$]. They included dryness of mouth [3 patients] »

10. P. Hanonnen 1998: A randomized double-blind placebo controlled study of moclobemide and amitriptyline in fibromyalgia

a) Caractéristiques de l'étude:

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles
Participants	N = 130 --> Groupe Amitriptyline N = 42, Groupe Moclobemide N = 43, Groupe Placebo N = 45. Femmes âgées de 18 à 65 ans, non hospitalisés, présentant les critères de la fibromyalgie selon la classification ACR de 1990 lors de la visite de dépistage et la visite initiale. Pour être inclus, les patients devaient avoir un score minimum de 4 (modéré) sur au moins trois des quatre échelles visuelles analogiques (EVA). Les items étaient : - (1) l'évaluation globale de l'état de santé général du patient (A quel point la Fibromyalgie a invalidé votre vie au cours de la semaine dernière ?) ; - (2) La douleur (" Quelle douleur avez-vous ressentie à cause de la Fibromyalgie au cours de la dernière semaine ?") ; - (3) la qualité et la quantité du sommeil (" A quel point les problèmes de sommeil ont perturbé votre vie au cours de la dernière semaine ?") ; - (4) la fatigue (« À quel point avez-vous été fatigué au cours de la semaine dernière ?").
Interventions	Les participants ont reçu au hasard l'une des 3 molécules ci-dessous : - Amitriptyline au coucher (initialement 12.5 mg jusqu'à max 37.5 mg à la visite de la semaine 6) + un Placebo matin et soir - Moclobemide matin et soir (initialement 150 mg jusqu'à max 600 mg à la visite de la semaine 6) + un Placebo au coucher - Placebo matin, soir et au coucher
Outcomes	<u>Objectif primaire :</u> - L'impression clinique de changement "The Clinical impression of the Change (CIC)" : mesurée par le médecin sur une échelle à 7 points. - L'amélioration moyenne du CIC des patients était également calculée. <u>Objectifs secondaires :</u> - La santé générale (General Health), la qualité et la quantité de sommeil, la douleur et la fatigue étaient évaluées sur une échelle visuelle analogique. - L'échelle fonctionnelle de Sheehan évaluant la gêne au travail, dans la vie sociale, les activités de loisirs et les devoirs de la vie familiale. - Le nombre de points sensibles "Tender points number" (jusqu'à 18) était réalisé à la visite de base, à 12 semaines ou au moment du départ si le patient abandonnait l'essai. - Le Nottingham Health Profile (NHP), auto-questionnaire sur la qualité de vie explorant 6 dimensions, était réalisé à la visite de base, à 12 semaines ou au moment du départ si le patient abandonnait l'essai. - L'impression clinique de la gravité de la fibromyalgie par le médecin (CIS), était réalisé à la visite de base, à 12 semaines ou au moment du départ si le patient abandonnait l'essai

	- Les effets secondaires, les maladies intercurrentes, la consommation médicamenteuse étaient enregistrées. - L'impression clinique globale de la Tolérance par le médecin (CGI) était utilisée comme mesure de la tolérance globale du médicament de l'essai.
Notes	Critères d'exclusion : Les personnes étaient exclues si elles souffraient de graves maladies cardiovasculaires, pulmonaires, hépatiques, hématologiques ou d'une néphropathie, d'un glaucome. Les femmes enceintes ou allaitantes, les personnes qui n'étaient pas disposés à interrompre tous les médicaments agissant sur le système nerveux central, les AINS et les analgésiques (autres que la paracétamol). Un dosage de la fonction thyroïdienne, du calcium sérique et de la créatine kinase devait également se situer dans les valeurs de référence. Pour exclure les patients souffrant de dépression majeure, de psychose, de trouble obsessionnel-compulsif ou d'un trouble lié à l'usage de drogues psychoactives, y compris la consommation d'alcool, présentement ou au cours des 6 derniers mois, un entretien clinique structuré (Scid-Ro) a été réalisée par un psychiatre avant l'inclusion. Étude multicentrique : 4 centres Durée 12 semaines Analyse en ITT Patients analysés N= 92 (A n= 32, P n= 30, M n= 30) Perdus de vue / exclus N= 40 perdus de vue Résultats significatifs pour l'Amitriptyline versus Placebo pour le CIC ainsi que pour l'amélioration moyenne du CIC. L'index du sommeil était significativement plus faible pour le groupe Amitriptyline versus Placebo à la fin de l'essai traduisant un meilleur sommeil. Le recours au paracétamol était moins fréquent significativement pour le groupe Amitriptyline versus Placebo.

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?		NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	« <i>The randomization was organised centrally with sequentially numbered envelopes consisting of blocks of six</i> »	PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« <i>Patients demographics did not differ between the treatment groups</i> » « <i>No statistically significant differences between the study treatment groups existed either in the number of tender points, CIS, the dimensions of NHP or the areas of sheehan's functional scales, or in the symptomatic VAS baseline</i> »	PN
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	« <i>the placebo capsules were identical to active drugs</i> »	NI
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?	« <i>The respective mean CGI of tolerabilities [...] (not significant)</i> » et CGI = Score de tolérabilité. " <i>The most typical adverse events, with at least a possible cause relationship to medication were headache and difficulties in falling asleep with MOCLO, dry mouth and fatigue with AMI and fatigue and headache with Placebo</i> " Cependant absence d'information donnée sur les pourcentages d'effets secondaires spécifiques de l'Amitriptyline comparé au Moclobemide et au Placebo.	NI
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITT selon les auteurs : « All the tests were two-sided and the intention-to-treat principle was applied » « since all patients received study medication, they are also included in the analyses. »	PY
2.7 Si N, PN / NI à 2.6 : Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		
Evaluation du risque de biais		Risque faible

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	38 perdus de vue au cours de l'essai / 130 patients randomisés soit 29% de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Probable utilisation de la méthode du dernier résultat rapporté dans la gestion des données manquantes : « <i>at the baseline and 12-week visits, or if the patient discontinued in the trial, the number of ACR [...] were recorded.</i> »	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Données manquantes dépendantes de la vraie valeur : abandon pour inefficacité ou effets secondaires : 29 patients / 130 soit 22% et motif de sortie d'étude non fourni pour les 9 derniers perdus de vue.	PY
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	« <i>Although the overall discontinuation rate was lowest in the AMI group, the differences in the overall rates or reasons for withdrawals between the treatment groups were statistically not significant</i> » mais cependant 29 % de perdus de vue	NI
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Évaluateurs : médecin et/ou patient Doute sur maintien du double aveugle du fait d'absence d'informations supplémentaire.	NI
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critère de jugement subjectif Évaluateur : patient et/ou médecin	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Protocole a priori non enregistré dans une base d'enregistrement	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Randomisation centralisée avec dissimulation de la séquence d'allocation Caractéristiques de la population initiale ne suggérant pas de problème de randomisation
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque faible	Absence d'informations pour statuer sur le double aveugle (Effets indésirables spécifiques de l'amitriptyline) A priori pas d'écarts par rapport à aux interventions prévues Analyse en ITT
Biais dû aux données manquantes	Risque élevé	29 % de données manquantes A priori méthode de gestion des données manquantes par “dernière visite rapportée” Données manquantes possiblement en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Doutes sur le maintien du double aveugle Critères de jugements subjectifs Evaluateur : patient et/ou médecin
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : Protocole non enregistré préalablement Absence de sélection de résultat ou d'analyse
Autres biais		Financement par le laboratoire Roche Oy Finland Multiplication des critères de jugements mais ajustement par la méthode de Bonferroni
Risque de biais global	Risque élevé	

d) Tableau des résultats :

Auteur : Hannonen Année : 1998					
Essai randomisé en parallèle 3 bras d'intervention Amitriptyline vs Moclobemide vs Placebo Etude en ITT, Participants N = 130 patients randomisés Perdus de vue/ exclus : 38					
	Critère de jugement	Mesure du critère (Avec valeur rapportée)	Amitriptyline	Placebo	Significatif ou non AT vs placebo
Critère de jugement primaire	Impression clinique de changement dans la sévérité de la fibromyalgie par le médecin (CIC)	Echelle -3 à +3 => positive si > ou égal 1 (Moyenne avec écart type)	1.14 [1.05]	0.38 [1.42]	p < 0.003 significatif
Critère de jugement secondaire	Santé générale	EVA 10 cm (AutoQ) (Moyenne et écart type)	4.4 [2.6]	5.3 [2.5]	Pas d'info
	Index de sommeil		4.5 [2.8]	5.2 [2.7]	P < 0.01 significatif
	Douleur		3.6 [2.8]	4.8 [2.9]	Pas d'info
	Fatigue		4.7 [2.8]	4.6 [2.6]	Pas d'info
	Mobilité	Nottingham Health Profile (autoQ) (Moyenne et écart type)	20.7 [18.3]	25.4 [17.7]	Pas d'info
	Energie		24.3 [27.5]	35.4 [35.8]	
	Douleur		36.6 [30.0]	49.3 [32.4]	
	Emotions		2.9 [7.7]	13.2 [20.2]	
	Sommeil		18.8 [25.3]	32.3 [30.3]	
	Vie sociale		3.3 [10.8]	1.6 [5.1]	
	Travail		4.5 [2.5]	5.3 [2.7]	
	Vie sociale		3.9 [2.6]	4.8 [2.7]	Pas d'info

	Devoirs de la vie familiale	Sheehan's functional scale (AutoQ) (Moyenne et écart type)	4.1 [2.5]	5.0 [2.7]	
	Impression clinique de sévérité par le médecin	Echelle de 1 à 7 avec : 1 = normal 7 = sévérité maximale (Moyenne et écart type)	3.4 [0.94]	3.78 [1.15]	Non significatif
	Points sensibles	Nb de points sensibles (sur 18) (Moyenne et écart type)	13.9 [4.2]	14.6 [3.6]	
	Effets indésirables	Nombre de patient avec ≥ 1 EI par groupe de traitement (en %)	74%	80%	Non significatif (p non donné)
		Relation causale supposée par le médecin entre EI et ttt reçu (en %)	43%	53%	
	Impression de tolérance par le médecin	Echelle de 1 à 4 : 1 = très faible 4 = très bonne (Moyenne et écart type)	2.9 [1.05]	2.64 [1.07]	Non significatif (p non donné)
	Recours à un médicament de secours (500mg paracétamol)	Nombre de comprimés pris durant les 84 jours d'essai (Moyenne et écart type)	40.0 [33.6]	73.1 [53.8]	p < 0.012 significatif
	Compliance	Différence entre le nombre de comprimés donnés et ceux rendus (En pourcentage)	97%	97%	Pas d'info

11. R.E Heymann 2001: A double-blind, randomized, controlled study of amitriptyline, nortriptyline and placebo in patients with fibromyalgia. An analysis of outcome measures.

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles.
Participants	N = 118, femmes âgées de 18 à 76 ans, non hospitalisés, présentant les critères de la fibromyalgie selon la classification ACR de 1990. Les participants devaient avoir lu et signé le consentement écrit informatif. Ils ne devaient pas avoir participé à une autre étude utilisant ces médicaments dans les 6 derniers mois, et devaient avoir arrêté tous les médicaments tels que les anxiolytiques, les neuroleptiques et/ou les antidépresseurs, les analgésiques, les narcotiques au moins 4 semaines avant le début de l'essai. L'utilisation de paracétamol était autorisée. Les patients ne devaient pas avoir utilisé l'amitriptyline ou la nortriptyline dans le passé ; ni souffrir d'arythmie cardiaque, d'insuffisance cardiaque, rénale ou hépatique, de glaucome, de rétention urinaire, d'hyperthyroïdie ou de maladies chroniques inflammatoires.
Interventions	Les participants ont reçu au hasard une des 3 molécules ci-dessous : - Un comprimé de 25 mg d'Amitriptyline au coucher - Un comprimé de 25 mg de Nortriptyline au coucher - Un comprimé de Placebo au coucher
Outcomes	<u>Critères de jugement principal :</u> - L'Evaluation de la qualité de vie avec le " Fibromyalgia Impact questionnaire " (FIQ) semble être le critère de jugement principal car il est utilisé pour le calcul du nombre de sujets nécessaires. <u>Critères de jugement secondaires :</u> - Le nombre de points sensibles "Tender points number" - L'auto-évaluation des patients de leur bien être par le "verbal scale global évaluation " (VSGI) - La compliance - Les effets indésirables
Notes	Critères d'exclusion : l'utilisation concomitante d'anxiolytiques, de neuroleptiques ou d'antidépresseurs ; la grossesse ; et le début d'une rééducation physique à visée thérapeutique (kinésithérapie, aquagym, etc.) concomitante au début de l'essai. Étude monocentrique Durée 8 semaines Analyse en ITTm Participants analysés N = 106 (A = 37, N = 36, P = 33) Perdus de vue N = 12 (3 Amitriptyline 2 Nortriptyline 7 Placebo), exclus N= 0 Résultat significatif dans l'amélioration sur le VGSI pour le groupe amitriptyline vs placebo.

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« amitriptyline, nortriptyline and placebo were all prepared by Sandoz Laboratory in standard packages for all groups, thus guaranteeing that the study remained blinded both investigator and patients. The randomization tables were supplied by the same laboratory. »	PY
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« Mean variance analysis (ANOVA) demonstrated no significant differences regarding age ($p=0.1413$), FIQ score ($p=0.5182$) and the number of tender points ($p=0.9909$), conforming the homogeneity of the three treatment groups »	PN
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur était assignée pendant l'essai ?	« <i>the most frequent side effects in the three groups the high incidence of side effects in the placebo group being the most remarkable fact. No significant differences were found among the three study groups ($\chi^2 = 4.75, p = 0.092$)</i> » De plus effets secondaires très similaires pour l'amitriptyline et la nortriptyline.	PN
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		PN
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné

2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à	ITTm avec 10,1 % de donnée manquantes	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	Les auteurs indiquent avoir utilisé un test Chi2 pour ajuster résultats aux perdus de vu, ceux-ci étaient principalement groupe placebo	NI
Evaluation du risque de biais		Risque Elevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	12 perdus de vue/118 => 10.1% de données manquantes	PN
3.2 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	« the Khi^2 to evaluate whether the number of drop-outs and the number of side effects in each group influenced the results » “the number of drop outs was smaller than those mentionned in other therapeutical studies on fibromyalgia and did not influence the final result of the study , probably due to the short treatment period.” Concernant les perdus de vue : « No significant differences among study groups were found ($Khi^2 = 3.500, p = 0.1737$) »	PY
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ? (		Non concerné
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?		Non concerné
Evaluation du risque de biais		Risque faible

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?	Même méthode de mesure pour les 3 groupes	PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Evaluation par patient principalement et tender points par examinateurs Double aveugle respecté	PN
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?		
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?		
Evaluation du risque de biais		Risque faible

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Protocole a priori non enregistré dans une base d'enregistrement	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	Une seule mesure de résultat	PN
5.3 ... Plusieurs analyses éligibles des données.	Non une seule analyse par donnée	PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Séquence de randomisation aléatoire et randomisation centralisée Les données démographiques de la population initiale ne suggèrent pas de problème de randomisation
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Maintien du double aveugle Absence d'écart à l'intervention Analyse en ITTm avec plus de 5% de perdus de vue Difficile de statuer sur un impact substantiel
Biais dû aux données manquantes	Risque faible	10.1 % de données manquantes sur les résultats Preuves indiquant qu'à priori pas de répercussion sur le résultat
Biais dû à la mesure de l'outcome	Risque faible	Méthode de mesure adaptée Maintien du double aveugle
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement Absence de sélection de résultats ou d'analyses
Autres biais		Pas d'informations sur les conflits d'intérêts Pas de méthode d'ajustement du risque alpha rapporté
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Heymann année : 2001					
Etude randomisée parallèle en 3 bras d'intervention : amitriptyline, nortriptyline, placebo					
Participants N = 118 patients (Amitriptyline (A) n = 40, Nortriptyline (N) n = 38, placebo (P) n = 40)					
Analyse en per protocole					
	Critères de jugement	Mesure du critère	Valeur rapportée	amitriptyline	Significatif ou non vs placebo
Critère de jugement primaire	Efficacité de l'amitriptyline et de la nortriptyline	FIQ	Moyenne et écart type	A: 39.97 ± 6.54 P: 51.68 ± 7.98 (36% d'amélioration par rapport baseline mais non significatif par rapport placebo : 24% amélioration)	Non significatif ([F(2,103) = 1.392 ; p = 0.253])
Critères de jugement secondaires	Efficacité et tolérance de l'amitriptyline et de la nortriptyline	Nombre de points sensibles	Moyenne et écart type	A: 14.2 ± 0.7 P: 14.7 ± 0.6 (Diminution de 13.9% groupe A vs 8.6% groupe P, non significatif)	Non significatif ([f(2,103) = 1.781 ; p = 0.173])
		Echelle verbale d'évaluation globale VSGI (amélioration)	Fréquence de distribution des patients %	31 patients (86.5%) groupe Amitriptyline vs 18 patients (54.5%) groupe Placebo	p = 0.00981
Autres		Effets indésirables		- Amitriptyline n=16 / 40(40%) (sécheresse buccale :4) - Nortriptyline n=31 / 38 (sécheresse buccale : 6) - Placebo n=25/40 (62.5%) (sécheresse buccale : 1)	p = 0.092 Non significatif
		Fréquence d'abandons		- AT 3 :7.5% (2 inconnus : 5% ; 1 refus de traitement 2.5%) - Placebo 7 (5 inconnus 12.5%, 1 arythmie ventriculaire 2.5%, 1 palpitation dyspnée 2.5%)	p = 0.1737 : Non significatif
		Compliance	Retour des flacons à la fin de l'étude	Pas de résultat	Pas d'info

12. E.A Fors 2002: The effect of guided imagery and amitriptyline on daily fibromyalgia pain: a prospective, randomized, controlled trial.

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles
Participants	N = 58 Femmes âgées de 21 à 68 ans, présentant les critères de la fibromyalgie selon la classification ACR de 1990.
Interventions	Les participants ont tout d'abord choisi de manière aveugle un des trois groupes : - Le groupe recevant une cassette dite " d'imagerie plaisante guidée " (Pleasant Guided Imagery) - Le groupe recevant une cassette dite " d'imagerie de l'attention" (The Attention Imagery) - Le groupe recevant une cassette vierge sans instruction Puis chaque participant a choisi aveuglément parmi les 2 traitements : - Un comprimé d'Amitriptyline débutée à 10 mg et augmentée jusqu'à la posologie de 50 mg à J11 - Un Placebo.
Outcomes	<u>Critères de jugement primaires :</u> Comparer les effets de deux différents types d'imagerie guidée sur la douleur Comparer l'effet de l'amitriptyline vs placebo sur la douleur <u>Critères de jugement secondaires :</u> Examiner les effets de l'anxiété et sur la douleur Mesures des critères de jugement : - La douleur notée sur EVA de 0 à 100 dans un journal quotidien. - Le STAI-T " State Trait Anxiety Inventory Trait" permettant d'évaluer l'anxiété. - LE BDI "Back Depression Inventory" permettant d'évaluer plusieurs aspects de la dépression. - Le ATQ-30 " Automatix negative Thoughts Questionnaire" permettant d'évaluer la fréquence des pensées négative associées à la dépression. - Les effets secondaires
Notes	Critères d'exclusions : sexe masculin Étude monocentrique, durée 4 semaines Analyse en ITTm Participants analysés N =55, Perdus de vue N = 3 Résultats non significatifs pour l'Amitriptyline versus Placebo

b) Tableau RoB2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	« Thereafter they were randomly assigned to one of the three psychological conditions by self-selecting an auditory cassette placed in an unmarked envelope. [...] They then blindly chose either amitriptyline or placebo, which had been repared in identical capsules by the hospital pharmacist »	Y
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?		Y
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« No significant baseline differences between the groups » “There was no difference between either the two randomized drug conditions or the three psychological interventions (PI, AI, CG) in the pain (VAS) pre-scores, age, education, symptom duration, depression (BDI), negative thoughts (ATQ) and trait-anxiety (STAI-T) of the participants (Table 1)” Mais pas de vérification au niveau du groupe amitriptyline et placebo qu'il y est bien une allocation 1 : 1 : 1 pour AI, PI et CG. (Or les auteurs constatent une différence d'efficacité entre PI d'une part et AI et CG d'autre part. Donc bien que le processus d'allocation soit correct, on ne peut pas être à l'abri d'une différence significative entre les groupes d'allocation due au hasard qui aurait un impact sur la vraie valeur d'AT vs placebo sur la douleur)	NI
Evaluation du risque de biais		Risque Faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Pas d'informations concernant les effets indésirables spécifiques à l'amitriptyline.	NI
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?	Double aveugle levé dans l'intervention en imagerie guidée.	NI
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITTm avec 5.17% de perdu de vue On considère d'après les informations fournies dans l'étude que 5 patients ont été exclus avant la randomisation et qu'il y a 3 perdus de vue par la suite.	PY
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?		
Evaluation du risque de biais		Risque faible

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient – elles disponibles pour tout ou presque tous les participants randomisés ?	On considère d'après les informations fournies dans l'étude que 5 patients ont été exclus avant la randomisation et qu'il y a 3 perdus de vue par la suite - 3 drop out / 58 patients => 5.17%	PY
3.2 SI N/PN/NI à la 3.1 Existe il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?		Non concerné
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?		Non concerné
3.4 Si Y/ PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?		Non concerné
Evaluation du risque de biais		Risque faible

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?	VAS : outil reconnu et outil d'analyse adapté SAS Proc Mixed (SAS 1992) « flexible ad powerful statistical procedure [...]it permits the specification of various error structures appropriate for repeated measures. »	PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?	Même outil de mesure	PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Doute sur le maintien du double aveugle	NI
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la	Critère de jugement subjectif Evaluateur : patient et/ou médecin	PY

connaissance de l'intervention reçue ?		
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?	1 seule mesure par résultats	PN
5.3 ... Plusieurs analyses éligibles des données.	Analyse en ITTm et per protocoles toutes les deux présentées.	PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Séquence d'allocation aléatoire et dissimulation de celle-ci Caractéristiques de la population initiale non fournies en sous-groupes
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque faible	Doutes concernant le maintien du double aveugle (absence d'informations concernant les effets secondaires spécifiques de l'amitriptyline) Analyse en ITTm avec environ 5% de données manquantes considérée comme appropriée
Biais dû aux données manquantes	Risque faible	5.17 % de données manquantes
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Doutes sur le maintien du double aveugle Critères de jugements subjectifs Evaluateur : patient et/ou médecin
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : Protocole non enregistré préalablement 2 Analyses réalisées et toutes deux présentées
Autres biais		Risque de biais de confusion devant absence de critères d'exclusions hormis le sexe masculin Absence de conflits d'intérêts Pas de méthode d'ajustement du risque alpha mentionnée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : Fors année : 2002						
Essai randomisé en parallèle 6 bras d'études : - " imagerie plaisante guidée " + Amitriptyline - " imagerie plaisante guidée " + Placebo - " imagerie de l'attention" + Amitriptyline - " imagerie de l'attention" + Placebo - Cassette vierge + Amitriptyline - Cassette vierge + Placebo Analyse en ITTm Nombre de participants N= 58						
	Critère de jugement	Mesure du critère	Analyse utilisée	Amitriptyline	Imagerie de l'attention (AI)	Paysage paisible et magnifique (PI)
Critère de jugement primaire	Douleur	Courbes de Moyenne et écart type traités par SAS Proc Mixed 1992	ITTm inapproprié	F (1.1452) = 0.00 Pas de différence significative par rapport au placebo p = 0.98	(t = 0.80, df = 1455) pas de différence significative par rapport groupe contrôle p > 0.05	(t = 2.93, df = 1455) Amélioration significative par rapport groupe contrôle p < 0.005
			Per protocol	F (2, 1453) = 0.16 Pas de différence significative par rapport placebo p = 0.85	Non fait	Non fait
			ITTm inapproprié	F (2.1452) = 0.28 p = 0.76 Pas d'interaction significative entre traitement psychologique et traitement par amitriptyline		
	Interaction entre Amitriptyline et traitement psychologique sur la douleur					
Critère de jugement secondaire	STAI-T	univariée	F (1.47) = 12.1, p = 0.001			
		multivarié	F (3.45) = 5.85, p = 0.002			
	BDI		F (1.47) = 14.4, p = 0.0001			

	ATQ			$F(1.47) = 5.4, p = 0.025$
	Ecart de mesure de douleur au premier jour de l'étude entre la mesure au laboratoire et celle dans le journal de recueil pour l'ensemble des participants	VAS de 0 à 10 cm Moyenne	Per protocol	7.2 paired t-test, $t = 2.36, df = 54, p = 0.022$

Critère de jugement :	Mesure du critère (Mean, SD)	Date	Amitriptyline non compliant n= 6	Amitriptyline compliant N = 22	Placebo N= 27	Significatif par rapport au placebo
Douleur sur carnet de bord	EVA (0 -> 100)	J1	39.0 (26.9)	42.6 (26.0)	44.7 (23.5)	Non significatif
		J7	30.1 (21.6)	38.3 (24.6)	41.9 (21.6)	
		J14	36.2 (35.1)	41.8 (27.8)	43.2 (21.9)	
		J21	33.1 (33.3)	34.1 (24.0)	49.2 (21.0)	
		J28	41.9 (36.2)	39.2 (29.1)	50.7 (24.1)	

13. A. Braz 2012: Effects of Panax ginseng extract in patients with fibromyalgia: a 12-week, randomized, double-blind, placebo-controlled trial.

a) Caractéristiques de l'étude

Méthode	Etude prospective, randomisée, en double aveugle, contrôlée par placebo, en groupes parallèles
Participants	Nombre de participants N = 52 Groupe Amitriptyline N = 16, Groupe Placebo N = 17, Groupe P. Ginseng N = 19. Femmes âgées de 21 à 61 ans, de toute origine ethnique, ayant un diagnostic de fibromyalgie remplissant les critères ACR 1990 définis par le Collège Américain de Rhumatologie, suivies en ambulatoire à la clinique de rhumatologie de Santa Rita, et ayant des résultats de biologies normales (NFS, Créatinine, transaminases, TSH).
Interventions	Les participants ont reçu soit : - De l' Amitriptyline hydrochloride à la dose de 25 mg par jour Soit - De l'extrait de racine de P. Ginseng à la dose de 100 mg par jour Soit - Un Placebo
Outcomes	Absence de distinction en critères de jugement primaire et secondaire : « The objective of the present study is to evaluate the therapeutic efficacy of P. ginseng root extract on the control of symptoms in patients with fibromyalgia » Mesure des critères de jugement : - Douleur évaluée sur une EVA de 0 à 10 - Fatigue évaluée sur une EVA de 0 à 10 - Qualité du sommeil évaluée sur une EVA de 0 à 10 - Anxiété évaluée sur une EVA de 0 à 10 - Mesure des points douloureux de la fibromyalgie “ Tender points” par une palpation digitale avec une pression d'approximativement 4kgf/cm² - La qualité de vie des patients était évaluée par le questionnaire FIQ (Fibromyalgia Impact Questionnaire)
Notes	Critères d'exclusion : Etaient exclus les patients présentant une maladie rhumatismale non traitée ou endocrinienne ou neurologique, une maladie rénale,

infectieuse ou osseuse ; Les patients atteints de glaucome, de rétention urinaire, de maladie coronarienne, d'arythmies ou d'insuffisance cardiaque congestive, les femmes enceintes et les femmes allaitantes.

Les patients qui avaient utilisé des antidépresseurs tricycliques au cours des trois mois précédents ; les patients hypersensibles au ginseng ou à son composant, les patients présentant des hémorragies, des taux élevés d'œstrogènes ou subissant un traitement hormonal ; les patients prenant des stéroïdes ou des médicaments antipsychotiques contre-indiqués avec l'utilisation de ginseng.

Avant le début de l'étude, les patients qui utilisaient des analgésiques, des médicaments opioïdes ou des anti inflammatoires devaient les avoir arrêtés leur pendant au moins 3 semaines avant le début de l'essai.

Etude monocentrique

Durée de l'essai : 12 semaines

Analyse en per protocole

Participants analysés N = 38 (Amitriptyline 13, Placebo 13 et P. Ginseng 12)

14 perdus de vue : N = 3 Amitriptyline, N = 4 Placebo et N = 7 P. Ginseng

Résultat significatif pour l'Amitriptyline versus Placebo pour l'anxiété à 6 semaines ($p < 0,05$).

b) Tableau RoB 2 :

Domaine 1 : Biais résultants du processus de randomisation		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
1.1 La séquence d'allocation est-elle aléatoire ?	Absence de d'informations supplémentaires sur la séquence d'allocation.	NI
1.2 La séquence d'allocation est-elle dissimulée jusqu'à inscription et répartition des participants aux interventions ?	Randomisation centralisée : « amitriptyline[...], P.ginseng root [...] and placebo [...] were prepared in identical form as capsules and provided to patients in sealed, black bottles. All were supplied by Orient Mix Fitoterapicos do Brasil Ltda »	PY
1.3 Les différences de bases entre les groupes d'intervention suggéraient-elles un problème au niveau du processus de randomisation ?	« There were no statistically significant differences between the three groups with respect to their baseline epidemiological or clinical characteristics (table 1) Table 1 : A n = 13 (16 à la randomisation), P n = 13 (17 à la randomisation), G n = 12 (19 à la randomisation), donc caractéristiques de bases des patients ayants terminés l'étude.	NI
Evaluation du risque de biais		Risque faible

Domaine 2 : Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention,)		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
2.1 Les participants étaient-ils au courant de l'intervention qui leur étaient assignée pendant l'essai ?	Pas d'information concernant les effets indésirables spécifiques de l'Amitriptyline	NI
2.2 Les soignants et les personnes dispensant les interventions étaient-ils au courant de l'intervention des patients pendant l'essai ?		NI
2.3 Si Y / PY / NI à 2.1 ou 2.2 : Y a-t-il eu des écarts par rapport à l'intervention envisagée en raison du contexte de l'essai ?		PN
2.4 Si Y / PY à 2.3 Ces écarts étaient-ils susceptibles d'avoir affecté les résultats ?		Non concerné
2.5 Si Y / PY / NI à 2.4 Ces écarts par rapport à l'intervention envisagée étaient-ils équilibrés entre les groupes ?		Non concerné
2.6 Une analyse appropriée a-t-elle été utilisée pour estimer l'effet de l'affectation à l'intervention	Analyse en ITTm : exclusion des données manquantes soit $14/52 = 26.9\%$, donc selon seuil probablement inappropriée	PN
2.7 Y a-t-il eu un impact substantiel (sur les participants) de l'échec de l'analyse des participants du groupe auquel ils ont été randomisés ?	Analyse inappropriée et 26.9% des données	PY
Evaluation du risque de biais		Risque élevé

Domaine 3 : Biais dû aux données manquantes		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
3.1 Les données des résultats étaient-elles disponibles pour tout ou presque tous les participants randomisés ?	14 patients perdus de vue sur 52 : 26.9% de données manquantes	PN
3.2 Existe-il des preuves que les résultats n'aient pas été biaisés par les données manquantes ?	Pas d'informations concernant la gestion des données manquantes	PN
3.3 Si N / PN à 3.2 Le manque de résultat pourrait-il dépendre de sa vraie valeur ?	Cause des perdus de vue : Effets secondaires : 8/14 Logistiques patients : 3/14 Pas d'information : 3/14 Plus de la moitié des abandons sont en lien avec la vraie valeur : 8/14, soit 8/52 => 15.4% des données manquantes sont en lien avec vraie valeur	PY
3.4 Si Y / PY / NI à 3.3 : Est-il probable que l'absence de résultat dépende de sa vraie valeur ?	26.9 % de données manquantes et lien avec la vraie valeur	PY
Evaluation du risque de biais		Risque élevé

Domaine 4 : Biais dû à des erreurs de mesure dans les résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
4.1 La méthode de mesure du résultat était-elle inappropriée ?		PN
4.2 La mesure ou la détermination du résultat aurait-elle pu différer d'un groupe d'intervention à l'autre ?		PN
4.3 Si N / PN / NI à 4.1 ou 4.2 : les évaluateurs des résultats étaient-ils au courant de l'intervention reçue par les participants à l'étude ?	Doute concernant le maintien du double aveugle	NI
4.4 Si Y / PY / NI à 4.3 : L'évaluation du résultat aurait-elle pu être influencée par la connaissance de l'intervention reçue ?	Critères de jugement subjectifs Evaluateurs = patient et/ ou médecin	PY
4.5 Si Y / PY / NI à 4.4 : Est-il probable que l'évaluation des résultats ait été influencée par la connaissance de l'intervention reçue ?	Pas assez d'éléments pour statuer	NI
Evaluation du risque de biais		Risque élevé

Domaine 5 : Biais dans la sélection des résultats		
Questions	Commentaires	Réponses (Y, PY, PN, N, NI, NA)
5.1 Les données qui ont produits ce résultat ont-elles été analysées conformément à un plan d'analyse prédéfini qui a été finalisé avant que les données sur les résultats sans insu soient disponibles pour l'analyse ?	Pas d'information concernant l'enregistrement préalable du protocole	NI
Le résultat numérique évalué est-il susceptible d'avoir été sélectionné, sur la base de résultats, parmi ...		
5.2 ... Plusieurs mesures de résultats éligibles (échelles, définitions, points temporels) dans le domaine de résultats ?		PN
5.3 ... Plusieurs analyses éligibles des données.		PN
Evaluation du risque de biais		Quelques inquiétudes

c) Résumé des domaines et jugement global :

Biais	Evaluation du risque de biais	Argumentation
Biais résultants du processus de randomisation	Risque faible	Peu d'informations concernant la séquence de randomisation, mais dissimulation de la séquence d'allocation grâce à une randomisation centralisée Les données démographiques de la population initiale ne sont pas présentées
Biais dû aux écarts par rapport aux interventions prévues (effet de l'affectation à l'intervention)	Risque élevé	Doute sur le maintien du double aveugle (absence d'informations sur les effets indésirables spécifiques de l'amitriptyline) Analyse en ITTm inappropriée avec plus de 5% de perdus de vue
Biais dû aux données manquantes	Risque élevé	26.9 % de données manquantes sur les résultats Absence d'information concernant la gestion des données manquantes Données manquantes en lien avec la vraie valeur
Biais dû à la mesure de l'outcome	Risque élevé	Méthode de mesure adaptée Maintien du double aveugle non statué Critères de jugements subjectifs
Biais dû à la sélection des résultats rapportés	Quelques inquiétudes	Pas d'information sur les intentions d'analyse préalables à l'essai : absence de temporalité dans le protocole. Protocole non enregistré préalablement Absence de sélection dans les résultats ni dans les analyses
Autres biais		Pas d'informations sur les conflits d'intérêts Pas de méthode d'ajustement du risque alpha rapportée
Risque de biais global	Risque élevé	

d) Tableau de résultats :

Auteur : BRAZ Année : 2012					
Etude randomisée parallèle en double aveugle 3 bras de traitements : Amitriptyline A, Placebo P, P. Ginseng G Durée : 12 semaines Participants N = 52 (amitriptyline 16, placebo 17, ginseng 19) Analyse en ITTm Fin de traitement N= 38 (A : 13, P 13, ginseng 12)					
	Critère de jugement	Mesure du critère	Valeur rapportée	amitriptyline	Significatif ou non versus Placebo
Critères de jugement	Douleur	EVA (0 à 10)	Moyennes de scores	A: 4.8 [0.8] P: 3.9 [0.8] Evolution intragroupe : F(5) = 8.8 , avec 12°S ttt réduction 44.6% de la douleur) p<0.001	p > 0.5 Non significatif
	Fatigue			A: 4.1 [0.9] P: 3.7 [0.4] Evolution intragroupe : F(5) = 5.9 avec 12°S ttt réduction 38.7% de la fatigue) p<0.001	p > 0.5 Non significatif
	Anxiété			Evolution comparée au placebo : F (2) = 5.6 soit une amélioration de moyenne de -2.3 par rapport au placebo en faveur de l'A Evolution intragroupe : F(5) = 7.0 avec 12°S ttt réduction 49.3% de l'anxiété) p < 0.001	p < 0.5 Significatif
	Qualité du sommeil			A: 2.8 [0.9] P: 2.6 [0.8]	p > 0.5

				Evolution intragroupe : F(5) = 5.8 avec 12°S ttt réduction 54.6% de la qualité du sommeil) P<0.001	Non significatif
	Points sensibles de la fibromyalgie		Moyenne et écart type	Evolution comparée au placebo : F (2) = 0.3 Evolution intragroupe : F(5) = 11.1 avec 12°S ttt réduction 41.7% du nombre de tender points) p<0.001	p > 0.5 Non significatif
	Qualité de vie	FIQ	Moyenne et écart type	A: 32 [6] P: 32 [4] Evolution intragroupe : F(5) = 3.0 avec 12°S ttt réduction 40.9% de la qualité de vie) P<0.05	p > 0.5 Non significatif

CONCLUSION GENERALE

TITRE : Revue systématique de la littérature sur l'efficacité de l'amitriptyline dans les indications « céphalées de tension » et fibromyalgie ».

CONCLUSIONS :

Les céphalées de tension chroniques ou permanentes concernent 2 à 5 % de la population adulte. Elles sont responsables pour le patient d'un risque d'abus médicamenteux et pour la société d'un coût important. La physiopathologie des céphalées de tension chroniques n'est pas complètement élucidée, ce qui explique qu'elles soient encore difficiles à traiter.

La fibromyalgie touche 1 à 9% de la population et bien qu'elle soit un motif fréquent de consultations en médecine générale, la prise en charge de ce syndrome est rendue difficile par l'absence de traitement étiologique et la mauvaise réponse aux traitements antalgiques. Reconnue depuis 1990 par le collège américain de rhumatologie et depuis 1992 par l'OMS, sa physiopathologie reste cependant inconnue. L'hypothèse la plus probable est un dérèglement du système central de la douleur.

L'amitriptyline, découverte en 1960, est reconnue à l'échelle mondiale pour son efficacité. Elle possède des propriétés antidépessives, analgésiques, et sédatives. Elle est la molécule préconisée en première intention pour le traitement des céphalées de tension chroniques dans plusieurs pays, dont la France, où elle a l'autorisation de mise sur le marché (AMM) dans cette indication. Elle est par ailleurs couramment utilisée pour le traitement des symptômes de la fibromyalgie, bien qu'elle n'ait pas l'AMM dans cette indication. En consultant la littérature, nous nous sommes aperçues qu'il n'existait aucune méta-analyse traitant de l'utilisation de l'amitriptyline dans les céphalées de tension. Une revue systématique de la littérature datant de 2008 conclut à un niveau de preuve SORT de niveau A. Cette revue de la littérature inclut cependant en majorité des études qui ne comparent pas l'amitriptyline à un placebo. En ce qui concerne la fibromyalgie, la méta-analyse la plus récente, *Amitriptyline for fibromyalgia in adults*, Andrew Moore et al parue en 2015, s'oppose aux résultats d'une méta-analyse plus ancienne : *Comparative efficacy and acceptability of amitriptyline, duloxetine and milnacipran in fibromyalgia syndrome: a systematic review with meta-analysis*, Häuser W publiée en 2011. En effet, elle remet en question la place de l'amitriptyline comme traitement de référence dans la recherche et dans la pratique clinique. Elle conclut que, du fait du haut niveau de risque de biais présent dans l'ensemble des essais cliniques inclus, le niveau de preuve de ses résultats est très faible, probablement surestimé, bien que les résultats montrent une efficacité dans la prise en charge de la douleur sur un quart des patients.

Notre travail s'inscrit dans le projet REB dirigé par le Pr Boussageon, qui vise à vérifier le niveau de preuve de l'usage de chaque médicament prescrit en médecine générale. Ainsi, nous avons réalisé une méta-analyse vérifiant l'intérêt de l'amitriptyline dans la prise en charge de la fibromyalgie et des céphalées de tension, en analysant son effet sur des critères de jugement pertinents : amélioration globale ressentie par le patient, amélioration globale ressentie par le médecin, intensité de la douleur, sommeil et fatigue pour la fibromyalgie et pour les céphalées de tension en termes de durée, intensité et fréquence des céphalées.

Pour l'indication céphalées de tension, nous avons inclus 7 études, toutes randomisées, en double aveugle et contrôlées par un placebo. Toutes nos études présentent un risque de biais global élevé,

selon l'outil Cochrane risk-of-bias 2 (RoB2). L'hétérogénéité des unités de mesure, des échelles et des scores utilisés par les auteurs ne nous a pas permis de réaliser une analyse quantitative. Nous avons cependant réalisé une analyse qualitative, qui suggère une différence en faveur de l'amitriptyline pour la durée, l'intensité et la fréquence des céphalées, avec des résultats plus ou moins significatifs selon les études. Selon la méthode REB, nous avons conclu qu'il existe "un manque de preuve" et selon GRADE que le niveau de preuve est "très faible".

Pour l'indication fibromyalgie, nous avons inclus 12 études, toutes randomisées, en double aveugle et contrôlées par un placebo. Le risque de biais global a été estimé élevé pour toutes les études selon l'outil RoB2. Les résultats des analyses qualitatives étaient toutes en accord avec les résultats des analyses quantitatives.

L'amélioration globale évaluée par le patient a été étudiée selon deux aspects : le ressenti global du patient et les répercussions de la pathologie sur la qualité de vie. Pour le ressenti global du patient, l'analyse quantitative a inclus 4 études, soit 190 patients. La Standardized Mean Difference (SMD) est significativement en faveur de l'amitriptyline, avec une taille d'effet modérée et une hétérogénéité modérée à importante. Pour les répercussions de la pathologie sur la qualité de vie, l'analyse quantitative a inclus 3 études, soit 143 patients. La SMD est en faveur de l'amitriptyline mais elle n'est pas statistiquement significative. L'amélioration globale évaluée par le médecin a permis d'inclure 4 études, soit 188 patients. La SMD est significativement en faveur de l'amitriptyline, avec une taille d'effet modérée et une hétérogénéité modérée. En ce qui concerne l'intensité de la douleur, nous avons inclus 9 études, soit 459 patients au total. La SMD est significativement en faveur de l'amitriptyline, avec une taille d'effet faible et une hétérogénéité importante. Pour le sommeil, nous avons inclus 6 études, soit 237 patients. La SMD est significativement en faveur de l'amitriptyline, avec une taille d'effet modérée et une hétérogénéité faible. En ce qui concerne la fatigue, nous avons inclus 6 études, soit 237 patients au total. Le résultat est en faveur de l'amitriptyline, mais il n'est pas statistiquement significatif.

Nous avons également analysé les effets indésirables et les perdus de vue. Pour les effets indésirables, nous avons inclus 3 études, pour un total de 185 patients. Le risque ratio montre une diminution non statistiquement significative du nombre d'effets indésirables dans le groupe placebo par rapport au groupe traité par amitriptyline. En ce qui concerne les perdus de vue, nous avons analysé 10 études, soit 663 patients. Le risque ratio montre, de façon statistiquement significative, qu'il y a plus de perdus de vue dans le groupe placebo que dans le groupe amitriptyline, avec une taille d'effet modérée et une hétérogénéité très faible.

Selon la méthodologie REB, pour l'ensemble des critères de jugement analysés, il existe "un manque de preuve" d'efficacité et selon GRADE le niveau de preuve est "très faible".

Ainsi, notre revue systématique de la littérature remet en question la place de traitement de référence de l'amitriptyline pour le traitement des céphalées de tension, dans la pratique clinique et dans la recherche. En effet, au vu des données de la littérature actuellement disponibles, l'efficacité de l'amitriptyline dans le traitement des céphalées de tension est « non probante ». Ceci est à mettre en balance avec le fait que les médecins ont observé son efficacité sur de nombreux patients au cours des années et que les résultats des études incluses suggèrent une différence en faveur de l'amitriptyline pour la durée, l'intensité et la fréquence des céphalées.

Notre méta-analyse confirme le manque de preuves quant à l'efficacité réelle de l'amitriptyline dans le traitement de la fibromyalgie. Elle suggère cependant un bénéfice du traitement, principalement

sur les troubles du sommeil (effet estimé modéré), et dans une moindre mesure, dans la réduction de la douleur et l'amélioration globale du patient.

Les conclusions de notre étude s'expliquent par le fait que toutes les études sont à haut risque de biais selon l'outil RoB2. Or l'évaluation du risque de biais a une place décisive dans la méthodologie REB et dans la détermination du niveau de preuve GRADE.

Notre travail a plusieurs limites : nous avons un faible nombre d'études, toutes sont anciennes, avec de faibles effectifs et une méthodologie peu détaillée. Par ailleurs, l'amitriptyline est une molécule pourvoyeuse d'effets indésirables spécifiques, tels que la sécheresse buccale et la somnolence, rendant le maintien du double aveugle difficile lors des essais. Ces effets indésirables ne sont a priori pas responsables d'abandons plus importants, étant donné que le pourcentage de perdus de vue est plus important dans le groupe placebo.

Nous préconisons pour ces raisons la réalisation d'un nouvel essai clinique pour chacune des deux pathologies étudiées. Ce dernier devrait utiliser une méthode appropriée de gestion des données manquantes et comparer l'amitriptyline au placebo et à une autre molécule active ayant des effets indésirables spécifiques semblables, afin de maintenir le double aveugle.

Nous considérons que l'amitriptyline a toujours sa place dans la prise en charge des céphalées de tension et dans la fibromyalgie, mais en étant conscientes d'une part de l'absence de preuve de son efficacité, et d'autre part de son effet limité pour la fibromyalgie et non quantifiable dans la céphalée de tension. Nous rejoignons les recommandations de prise en charge actuelle pour la fibromyalgie, qui préconisent que la prise en charge des patients soit en priorité axée sur des thérapeutiques non pharmacologiques

VU ET PERMIS D'IMPRIMER

Grenoble, le 01.06.2022

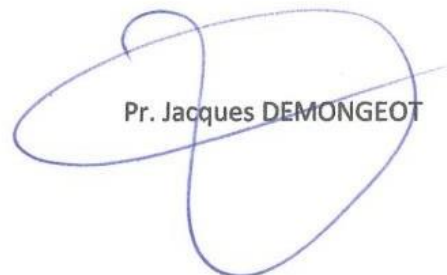
LE DOYEN DE L'UFR DE MÉDECINE



Pr. Patrice MORAND

Pour le Président
et par délégation
—
Le Doyen de Médecine
Pr. Patrice MORAND

LE PRÉSIDENT DU JURY :

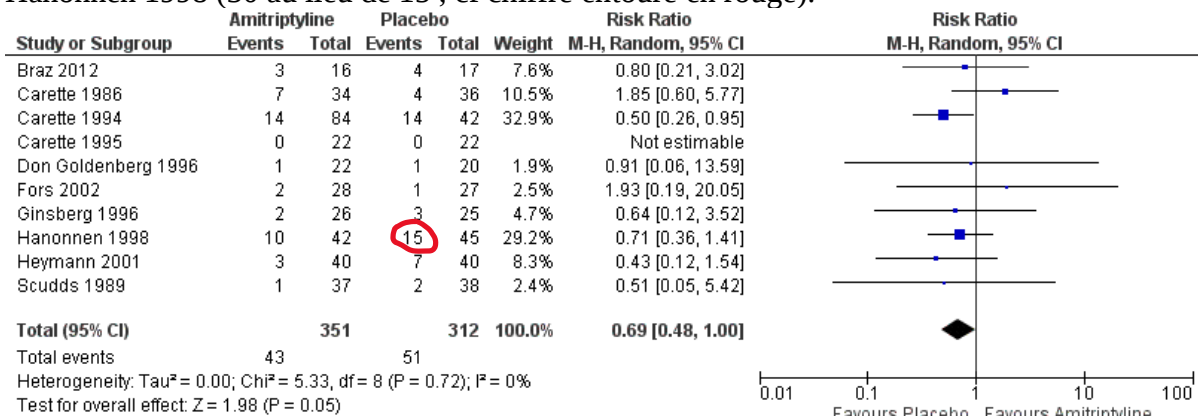


Pr. Jacques DEMONGEOT

Erratum

Erreur dans les conclusions de la thèse :

Nous avons corrigé avant la soutenance une erreur d'inattention concernant le critère de jugement « nombre de perdus de vue ». Nous avons inversé par inadvertance, le nombre de perdus de vue dans le groupe placebo et le nombre de patients ayant terminés l'étude dans ce même groupe, pour l'étude Hanonnen 1998 (30 au lieu de 15 , cf chiffre entouré en rouge).



Ceci a conduit à une modification du résultat pour le critère de jugement « nombre de perdus de vue » qui est devenu non significatif, bien qu'il soit toujours en faveur d'une diminution du nombre de cas dans le groupe amitriptyline. Ainsi, la conclusion signée par le doyen Monsieur le Professeur Pr Patrice MORAN, qui n'était plus modifiable lors de notre correction, contient une erreur. En effet, il est noté « En ce qui concerne les perdus de vue, nous avons analysé 10 études soit 663 patients. Le risque ratio montre, de façon **statistiquement significative**, qu'il y a plus de perdus de vue dans le groupe placebo que dans le groupe amitriptyline, avec une taille d'effet modérée et une hétérogénéité très faible. ». La correction des nombres de perdus de vue amène à conclure pour ce critère de jugement « En ce qui concerne les perdus de vue, nous avons analysé 10 études soit 663 patients. Le risque ratio suggère qu'il y ait plus de perdus de vue dans le groupe placebo que dans le groupe amitriptyline, mais ce de façon non significative. »