



HAUTE AUTORITÉ DE SANTÉ

RAPPORT D'ÉVALUATION TECHNOLOGIQUE

Appendicectomie

Éléments décisionnels pour une indication pertinente

Novembre 2012

Ce rapport d'évaluation technologique est téléchargeable sur
www.has-sante.fr

Haute Autorité de santé

Service documentation – Information des publics
2, avenue du Stade de France – F 93218 Saint-Denis La Plaine Cedex
Tél. : +33 (0)1 55 93 70 00 – Fax : +33 (0)1 55 93 74 00

Sommaire

Liste des abréviations	5
1. La demande.....	6
1.1 Contexte de la demande	6
1.2 Objectifs de la demande.....	6
2. Eléments de contexte	7
2.1 Pertinence de la réalisation des soins, principes généraux.....	7
2.2 Historique des référentiels français.....	8
2.3 Données de pratique d'appendicectomie en France	8
2.4 Généralités sur l'appendicite	9
2.5 Diagnostic de l'appendicite.....	13
3. Problématique et objectifs du travail.....	16
3.1 Définition de la problématique.....	16
3.2 Objectifs du travail	16
4. Méthode de travail	18
4.1 Analyse de la littérature	18
4.2 Position des professionnels de santé.....	20
5. Analyse de la littérature.....	24
5.1 Questions traitées et limites méthodologiques	24
5.2 Les signes cliniques et biologiques.....	25
5.3 Performance diagnostique de l'imagerie.....	43
5.4 La coelioscopie exploratrice.....	59
5.5 Le traitement antibiotique versus l'appendicectomie.....	63
6. Avis des professionnels.....	68
6.1 Concept de la pertinence	68
6.2 Evolution du nombre d'appendicectomies et variabilité interrégionale.....	68
6.3 Diagnostic clinique et biologique	69
6.4 Les examens d'imagerie.....	70
6.5 La coelioscopie exploratrice.....	71
6.6 Eléments décisionnels formalisés sous forme de propositions	72
6.7 Proposition d'un arbre décisionnel devant une suspicion d'appendicite	74
Conclusion générale.....	79
Annexe 1. Paramètres de l'évaluation diagnostique : quelques définitions.....	82
Annexe 2. Codes GHM utilisés pour la recherche des données de pratique	84
Annexe 3. La stratégie de recherche et la liste des sources interrogées	85
Annexe 4. Grille AMSTAR	90
Annexe 5. Compte rendu de la première réunion du groupe d'orientation du 15 décembre 2011	92

Annexe 6.	Compte rendu de l'audition du président de la SFMU du 06 janvier 2012	98
Annexe 7.	Compte rendu de la deuxième réunion du groupe d'orientation du 03 mai 2012.....	100
Annexe 8.	Niveau de preuve scientifique fourni par le grade des recommandations.....	107
Annexe 9.	Tableaux d'évidence.....	108
Annexe 10.	Analyse des réponses après cotation.....	132
Annexe 11.	Recueil de l'avis des professionnels sur la qualité de l'argumentaire scientifique	133
Annexe 12.	Liste des tableaux.....	134
Participants.....		135
Références		136
Fiche descriptive		140

Liste des abréviations

ACEP	<i>American College of Emergency Physicians</i>
ACOI	<i>Associazione dei Chirurghi Ospedalieri Italiani</i>
ACR	<i>American College of Radiology</i>
AHRQ	<i>Agency for Healthcare Research and Quality</i>
ANDEM	Agence nationale pour le développement de l'évaluation médicale
ATIH	Agence technique de l'information et de l'hospitalisation
AUC	Aire sous la courbe (<i>Area Under Curve</i>)
CNAMTS	Caisse nationale de l'assurance maladie des travailleurs salariés
CRP	<i>C-Reactive Protein</i>
DSS	Direction de la sécurité sociale
DGOS	Direction générale de l'offre de soins
DAA	Douleur abdominale aiguë
EAES	<i>European Association of Endoscopic Surgery</i>
FHF	Fédération hospitalière de France
FID	Fosse iliaque droite
GB	Globules blancs
GHM	Groupe homogène de malades
HAS	Haute autorité de santé
NFS	Numération formule sanguine
NIH	<i>American National Institute of Health</i>
RR	Risque relatif
SAGES	<i>Society of American Gastrointestinal and Endoscopic Surgeons</i>
SFCE	Société française de chirurgie endoscopique
SFCL	Société française de chirurgie par laparoscopie

1. La demande

1.1 Contexte de la demande

Cette demande a été émise en juin 2010 par le Ministère de la santé (DSS et DGOS) et par la CNAMTS en vue du programme 2011. Elle s'inscrit dans le cadre d'une démarche générale dont l'objectif est l'évaluation de la pertinence des actes réalisés en établissement de santé et la justification de leur prescription.

Cette saisine concerne une série d'actes, identifiés par la Fédération hospitalière de France (FHF) et qui interpellent par la fréquence et/ou par la variation de leur taux de recours, géographique, entre secteurs publics et privés, ou inter-établissements. Ces écarts de pratique soulèvent des interrogations sur le bien fondé de leur réalisation.

Dans ce contexte, la DGOS, la DSS et l'Assurance maladie souhaitent que soient produits, pour une liste d'actes et de façon pluriannuelle, des référentiels de bonne pratique clarifiant les indications médicales pertinentes pour ces actes et les séjours en établissements de santé qui leur sont associés. Parmi les actes retenus pour évaluation en 2011 figure l'appendicectomie.

En ce qui concerne l'appendicectomie, la FHF a réalisé des études sur la pratique de cet acte chirurgical en France (1). L'analyse de ces études et notamment de la répartition géographique de la fréquence de l'acte révèle des disparités importantes entre régions : plus fréquente dans des zones rurales et des villes de petites et moyennes tailles et inversement proportionnelle au dynamisme démographique et à la population du territoire. Les facteurs démographiques et épidémiologiques, ainsi que la difficulté d'assurer dans ces territoires une permanence de soins efficace, en particulier au niveau de la radiologie sont des hypothèses émises pour expliquer cette variabilité régionale. Cependant, ils ne seraient pas suffisants pour la justifier et une part de cette disparité pourrait être imputée à des variations de pratique médicale.

1.2 Objectifs de la demande

L'objectif du demandeur est de disposer des référentiels qui définissent les critères permettant de décider de la réalisation d'un acte donné. Ces référentiels seraient d'une part diffusés auprès des professionnels pour contribuer à l'amélioration de leur pratique et d'autre part, permettraient aux décideurs d'évaluer le bien fondé des actes réalisés et de prendre de mesures relatives à :

- l'optimisation de la prise en charge des patients par une organisation des soins efficace et rationalisée qui garantit la qualité et la sécurité des soins tout en limitant la réalisation d'actes inutiles ;
- la maîtrise médicalisée des actes avec une révision des modalités de prise en charge et de tarification ainsi que par une réduction des examens non pertinents.

En ce qui concerne l'appendicectomie, le référentiel sera focalisé sur la définition de la stratégie diagnostique permettant de poser l'indication de cet acte.

2. Eléments de contexte

2.1 Pertinence de la réalisation des soins, principes généraux

2.1.1 Définition de la notion de pertinence

La *RAND Corporation* (*University of California, Los Angeles*) propose la définition suivante de la pertinence : « *[Un traitement est approprié quand] le bénéfice escompté pour la santé (p. ex augmentation de l'espérance de vie, soulagement de la douleur, réduction de l'angoisse, amélioration de capacités fonctionnelles) est supérieur aux conséquences négatives attendues (p. ex mortalité, morbidité, anxiété, douleur, durée d'arrêt de travail) d'une façon suffisante pour estimer qu'il est valable d'entreprendre la procédure, indépendamment de son coût.* » (2).

Cette définition a été étendue par le NHS en intégrant des dimensions additionnelles : « *Nous estimons qu'il manque deux dimensions importantes à [la définition de la pertinence par la RAND Corporation] : l'individualité du patient concerné et la disponibilité des ressources de santé. Nous suggérons la définition suivante :*

Pertinent, pour un soin, signifie qu'il a été choisi parmi l'ensemble des interventions disponibles qui ont démontré leur efficacité pour une affection, comme étant le plus vraisemblablement à même de produire les résultats attendus pour un patient donné. Une intervention ne peut être pertinente que si certaines conditions sont satisfaites. Les compétences techniques et les autres ressources nécessaires à l'intervention doivent être disponibles, en sorte qu'il puisse être dispensé selon les bons standards. L'intervention doit être réalisée d'une manière telle qu'elle soit acceptable pour le patient. Les patients doivent recevoir une information adéquate au sujet de toutes les interventions potentiellement efficaces. Leurs préférences sont centrales dans la détermination de quelle intervention sera pertinente parmi celles dont l'efficacité est connue. Leurs préférences seront indicatives non seulement de l'objectif principal qu'ils espèrent atteindre, mais aussi de leurs perceptions des effets secondaires qui pourraient advenir. La pertinence des interventions de santé doit également être considérée dans le contexte social et culturel actuel, et au regard de la justice de la répartition des ressources de santé. » (3).

Sur la base des deux définitions précédentes, la HAS retient que l'analyse de la pertinence d'une intervention de santé comporte potentiellement plusieurs dimensions :

- la balance entre les bénéfices et les risques ;
- la probabilité pour l'intervention d'aboutir aux résultats attendus (par comparaison avec d'autres traitements) ;
- la qualité de l'intervention de santé (au regard de standards) ;
- la prise en compte des préférences des patients (ce qui implique une information appropriés) ;
- la prise en compte du contexte social, culturel et de la disponibilité des ressources de santé.

En fonction de la saisine et des données disponibles, les travaux portent sur une ou plusieurs des dimensions énoncées ci-dessus. Le périmètre du travail est, pour chaque saisine, défini lors d'une étape de cadrage.

Pour certaines saisines, la notion « d'intervention » devra être comprise dans un sens large renvoyant par exemple au parcours de soin du patient.

2.1.2 Elaboration de référentiels de pertinence

Différentes méthodes peuvent être utilisées pour élaborer des référentiels de pertinence. Elles comprennent dans tous les cas :

- une analyse de l'ensemble des données scientifiques spécifiques ;
- le recueil d'avis experts, pour lequel plusieurs méthodes formalisées existent.

La méthode choisie dépend fortement du degré de preuves disponibles dans la littérature.

Quelle que soit la méthode entreprise, la génération de référentiels de pertinence doit être précédée de l'identification des questions sur l'intervention de santé auxquels ils se rapportent, en prenant en compte les objectifs, les stratégies d'intervention et le contexte, ainsi que l'optimalité de la balance bénéfice-risque.

Tout référentiel de pertinence a, par construction, deux utilisations possibles découlant du fait qu'il énonce nécessairement des paramètres permettant de décrire des situations pour lesquelles on cherche à déterminer si une action est appropriée. La première est l'utilisation pour guider la décision clinique (sans se substituer à la démarche individuelle du praticien). La seconde est l'utilisation rétrospective dans une démarche d'évaluation des pratiques. La génération d'un référentiel de pertinence peut avoir comme objectif principal l'un ou l'autre de ces usages.

L'objectif du présent projet se place dans l'axe de la première de ces utilisations, la méthode de travail proposée est développée dans un chapitre ultérieur de ce document.

2.2 Historique des référentiels français

En 1996, l'ANDEM en collaboration avec les professionnels, a pu dégager un consensus sur les pratiques recommandables en matière de diagnostic et de prise en charge de l'appendicite. Les mêmes recommandations ont été publiées par la SNFGE en 2001. Depuis aucun référentiel français n'a été publié sur le sujet.

2.3 Données de pratique d'appendicectomie en France

La base nationale de l'Agence technique de l'information sur l'hospitalisation (ATIH) a été interrogée afin d'estimer le nombre d'actes réalisés en France entre 1999 et 2009 tous établissements confondus (publics et privés). Les codes GHM¹ utilisés pour cette recherche sont décrits en Annexe 2.

Sur les dix dernières années, le nombre total d'actes d'appendicectomies réalisés en France a connu une diminution progressive de l'ordre de 34% entre 1999 et 2009. En revanche, le nombre d'appendicectomies pour des appendicites compliquées a enregistré une progression médiane de 14% depuis 2005 (4).

L'Assurance maladie s'est intéressée à la diminution du nombre d'appendicectomies en fonction de l'âge et du sexe et constate sur les données recueillies entre 2006 et 2010 que la baisse concerne plus particulièrement les jeunes de moins de 20 ans et les femmes, avec un taux respectivement de (-20%) et (-14%).

En ce qui concerne la variabilité interrégionale évoquée par les demandeurs, les données régionales pour les années 2006 et 2008, tous établissements confondus publics et privés, ont été recherchées sur la base nationale de l'ATIH. Les données montrent sur chiffres bruts, qu'il peut exister un facteur de 1,5 à 2 entre les régions de faible incidence et les régions de forte incidence (4).

Lors de la séance² de l'académie nationale de chirurgie consacrée à l'appendicite, l'Assurance maladie a présenté les données d'une analyse interdépartementale pour identifier la variation géographique du taux de recours à l'appendicectomie/10000 habitants pour l'année 2010. Les chiffres ont été recueillis chez les patients d'un département donné sans tenir compte du département où a eu lieu l'intervention. Pour la France métropolitaine, l'Ile-de-France, apparait comme la région repère (taux le plus bas) avec 15,65 appendicectomies/10000 habitants (chez les moins de

¹ La classification en groupes homogènes de malade (GHM)

² Séance de l'académie nationale de chirurgie du 30 mai 2012

20 ans). Dans les DOM-TOM, Martinique, Guadeloupe et Guyane, les taux sont en dessous de 3 appendicectomies/10000 habitants.

Les facteurs qui ont été identifiés par la CNAMTS (après une analyse de régression) comme étant associés aux variations interdépartementales de recours aux soins, sont : l'offre de soins, la densité et croissance démographique des départements, les patients ou les types d'interventions.

- une corrélation positive est retrouvée avec l'âge et le sexe : plus il ya une population de moins de 20 ans et de femmes plus il ya recours aux interventions ;
- une corrélation négative est retrouvée entre le nombre d'appendicectomies et la spécialisation des pôles de chirurgie viscérale : plus les services pratiquent d'interventions, moins le taux d'appendicectomies est important par rapport à l'ensemble de leur activité de chirurgie digestive ;
- une corrélation négative est retrouvée entre la part de la durée des séjours de niveau 1 (PMSI) et le nombre d'interventions : plus la durée de séjour est courte et plus le nombre d'interventions est important dans le département.

Une approche par établissement a permis de ressortir les facteurs supplémentaires suivants :

- la part d'appendicectomies dans l'activité de la chirurgie viscérale, le chiffre moyen sur 566 établissements est de 50/an ; on observe que plus du tiers des établissements présentent un taux double du chiffre moyen ;
- le recours à l'exploration d'imagerie avant intervention (écho, scanner, IRM) : une différence importante est observée en fonction des établissements, le taux global est de 70% et le recours plus ou moins important selon les structures.

La CNAMTS considère, au regard des chiffres de recours aux soins, que la variation interdépartementales est confirmée, que certains facteurs explicatifs peuvent être avancés mais que ces facteurs sont loin de tout expliquer. La CNAMTS conclut en conséquence qu'il est nécessaire aujourd'hui d'analyser tout le parcours des soins.

2.4 Généralités sur l'appendicite

Depuis le début du XXI^{ème} siècle, l'incidence de l'appendicite est en baisse dans les pays industrialisés en raison de plusieurs facteurs notamment la modification des habitudes alimentaires avec un apport plus important de fibres et une amélioration de l'hygiène qui a réduit le nombre des infections entériques (5,6).

Néanmoins, l'appendicite aiguë reste encore aujourd'hui la cause la plus fréquente d'hospitalisation pour syndrome douloureux abdominal aiguë. Elle représente près de 30% des interventions de chirurgie digestive (7,8). L'incidence de l'appendicectomie en France en 2009 est de 1,39/1000 habitants.

La majorité des cas d'appendicite aiguë surviennent chez l'adolescent et le jeune adulte entre la 2^{ème} et la 3^{ème} décennie avec une prédominance masculine dans cette tranche d'âge. Elle est relativement rare chez l'enfant de moins de 3 ans et chez le sujet âgé. Cependant, c'est dans ces tranches d'âge que les formes compliquées (perforées) sont relativement plus fréquentes et que la mortalité est plus élevée. L'incidence de perforation est de 70 à 80% chez l'enfant et de 30% chez le sujet âgé (5).

Le risque principal d'une appendicite est l'évolution vers la perforation puis la péritonite généralisée qui met en jeu le pronostic vital : la mortalité est de 0,1% dans les formes non compliquées, elle est de 1,5 à 5% en cas de perforation appendiculaire (9). Le pronostic est directement lié à la précocité du diagnostic et du traitement.

Le traitement recommandé de l'appendicite aiguë est à ce jour chirurgical ; la seule question qui se pose est celle de l'urgence de l'intervention. Une appendicectomie peut être réalisée soit par laparotomie en fosse iliaque droite qui reste le traitement de référence, soit par cœlioscopie (10-13).

L'appendicectomie est un acte courant, il n'est néanmoins pas sans risques et peut entraîner des complications spécifiques (péritonite par lâchage du moignon, abcès postopératoires...) et des complications inhérentes à tout acte de chirurgie abdominale (phlébite, embolie pulmonaire, hémorragie, infection au niveau de l'incision, de la sonde urinaire, d'une voie veineuse, occlusion et bride intra-péritonéale...).

Ainsi l'exérèse d'un appendice sain n'est pas sans conséquences car elle ne met pas à l'abri des complications spécifiques.

Concernant le traitement conservateur par antibiothérapie exclusive, quelques études (14,15) comparant la chirurgie *versus* l'antibiothérapie ont montré un intérêt possible de cette dernière modalité de soins. Une récente étude française, randomisé multicentrique (16), réalisée par six services hospitaliers de l'APHP a conclu que l'appendicectomie en urgence devait rester le gold standard de la prise en charge des appendicites non compliquées.

Cependant, ces conclusions ont été remises en question en raison de l'inclusion à tort dans cet essai de certains patients avec appendicites compliquées et également l'utilisation d'une association d'antibiotiques susceptible de rencontrer une résistance de certains germes dans cette pathologie (17).

Le débat sur le traitement conservateur n'est donc pas clos et les récents éléments³ d'actualité ont imposé de se pencher sur cette question non prévue initialement dans ce travail compte-tenu de l'orientation exclusivement diagnostique qui a été validée pour cette évaluation (4). Un chapitre consacré à l'analyse comparative des deux stratégies thérapeutiques a donc été réalisé.

2.4.1 Physiopathologie

L'infection appendiculaire est provoquée par une obstruction de la lumière de l'appendice, par un obstacle (9) :

- endoluminal (coprolithe, corps étranger ingéré, ascaris, etc.) ;
- pariétal (hyperplasie lymphoïde, tumeur appendiculaire carcinoïde) ;
- caecal.

Cette obstruction provoque une pullulation microbienne dans la lumière appendiculaire ainsi que des lésions muqueuses puis pan pariétales associant un infiltrat inflammatoire à des lésions infectieuses et ischémiques (7).

La réaction inflammatoire peut aller de la simple congestion de l'appendice associé à une dilatation des vaisseaux de la séreuse (appendice inflammatoire), à une augmentation du volume de l'organe qui devient œdémateux et qui, à l'occasion d'une surinfection, se recouvre de fausses membranes (appendice suppuré) (7).

2.4.2 Anatomopathologie

L'examen anatomopathologique d'une pièce d'appendicectomie peut montrer des lésions plus ou moins sévères allant d'une simple infiltration muqueuse à une suppuration et une nécrose de toute la paroi (7).

Les divers aspects macroscopiques d'une appendicite sont décrits ci-dessous et correspondent à des stades de gravité croissante (11) :

- l'appendicite catarrhale : l'appendice est hyperhémie avec une accentuation de la vascularisation sous-séreuse ;
- l'appendicite ulcéreuse et suppurée : l'appendice est rouge, œdématisé, avec parfois quelques fausses membranes ;

³ Séance de l'académie de médecine consacrée à l'appendicite et articles parus dans la presse consécutivement à cette séance : l'APM le 30 05 2012, le quotidien du médecin le 31 05 2012, Panorama du médecin le 01 06 2012, le Figaro le 04 06 2012, le quotidien du pharmacien le 11 06 2012.

- l'appendicite abcédée : elle correspond à un appendice ulcéré et suppuré avec une inflammation péri-appendiculaire intense et une paroi infiltrée de micro-abcès ;
- l'appendicite phlegmoneuse : l'appendice devient turgescence se couvre de fausses membranes, sa lumière contient du pus et il se produit une nécrose suppurée de sa paroi ;
- l'appendicite gangréneuse : des plages nécrotiques apparaissent sur l'appendice, il s'installe une obstruction veineuse liée à l'œdème, parfois étendues au cæcum. Le risque de perforation est alors maximal.

Il n'y pas de parallélisme entre l'intensité des lésions anatomiques et la gravité clinique. Il est ainsi possible de découvrir lors de l'intervention des lésions très évoluées, pré-perforatives alors que les symptômes cliniques sont peu marqués (9).

2.4.3 Les différentes formes cliniques de l'appendicite

► Appendicite aiguë en position iliaque droite

Rencontrée dans environ 50 à 60% des cas, la forme dite « typique » est celle de l'appendicite aiguë en position iliaque droite (5).

Les signes fonctionnels sont une douleur de la fosse iliaque droite ou une douleur d'abord épigastrique puis localisée à droite, des vomissements, une constipation ou parfois une diarrhée. Les signes généraux sont une fièvre aux alentours de 38°C. Les signes physiques sont un abdomen plat et une douleur provoquée au point de Mac Burney⁴ avec une défense à la palpation. Il existe au toucher rectal une douleur à droite dans la moitié des cas. Sur le plan biologique, Il existe très souvent une hyperleucocytose à polynucléaires neutrophiles, elle est associée à une appendicite évoluant depuis plusieurs heures. Elle constitue un des arguments du faisceau diagnostique (18,19).

► Les autres formes cliniques

Dans environ 20 à 30% des cas, l'appendicite a une présentation clinique atypique (20). Les différentes formes cliniques de l'appendicite se déclinent selon différents aspects : la symptomatologie, la topographie et le terrain.

► Selon la symptomatologie

- **les formes frustes** : l'appendicite peut évoquer le tableau d'une gastrite, particulièrement chez le sujet jeune, où la symptomatologie se résume à des douleurs épigastriques et à des vomissements (21) ;
- **les formes évolutives** :
 - *le plastron appendiculaire* : qui associe un état infectieux persistant à des troubles du transit et à une masse mal limitée dans la fosse iliaque droite. L'évolution peut se faire vers la guérison ou vers l'abcès péri-appendiculaire,
 - *l'abcès d'emblée* : signes de suppuration profonde. Localement, il existe une masse de la fosse iliaque droite ferme, limitée, exquisément douloureuse. L'abcès peut se rompre dans la grande cavité péritonéale, provoquant une péritonite aiguë généralisée, ou fistuliser dans la vessie ou le colon,
 - *la péritonite généralisée* : d'emblée ou secondaire à la rupture d'un abcès péri-appendiculaire (19).

► Selon la topographie

- **appendicite pelvienne** : qui peut simuler une salpingite aiguë. Les douleurs, localisées à la région hypogastrique droite, sont au premier plan, associées le plus souvent à un fébricule,

⁴ Point de McBurney (M). Point situé au tiers externe de la ligne reliant l'épine iliaque antéro-supérieure droite à l'ombilic.

mais les signes fonctionnels et physiques sont unilatéraux. Il peut exister des signes fonctionnels urinaires (pollakiurie) trompeurs ;

- **appendicite sous-hépatique** : qui peut simuler une cholécystite aiguë : douleurs de l'hypochondre droit, fièvre et défense mais où l'échographie révèle une vésicule dont la paroi est fine et normale et élimine la cholécystite ;
- **appendicite rétro-caecale** : l'appendice est situé en arrière du cæcum et en avant du muscle psoas, les douleurs sont plutôt lombaires droites et le patient a parfois une position antalgique en psöitis (cuisse droite fléchie sur le tronc), il n'y a pas de signes fonctionnels urinaires ;
- **appendicite méso-cœliaque** : où l'appendice est situé au milieu des anses grêles. Le tableau clinique est celui d'une occlusion associée à de la fièvre qui impose une intervention en urgence ;
- **appendicite « à gauche »** : exceptionnellement, la symptomatologie peut être inversée en cas de *situs inversus* ;
- **appendicite herniaire** : dans certains cas l'appendice peut avoir migré dans une hernie inguinale droite et simuler un étranglement herniaire.
- (7,18,19).

► Selon le terrain

Les différentes formes cliniques selon le terrain se distinguent par la difficulté à poser un diagnostic rapide et par l'aggravation qui en résulte en termes de pronostic.

- **appendicite de la femme enceinte** : pendant la grossesse, la fréquence de l'appendicite n'est pas augmentée (0,5 à 1/1000 grossesses). Son diagnostic est néanmoins rendu difficile en raison de la modification de la position du cæcum liée à l'utérus gravide qui le refoule vers le haut. Le diagnostic différentiel avec une pyélonéphrite ou avec cholécystite aiguë peut être difficile. La gravité de l'appendicite chez la femme enceinte est liée au retard diagnostique et à une plus grande tendance à la diffusion péritonéale (18,21) ;
- **appendicite du sujet âgé** : chez la personne âgée, un retard diagnostique est fréquent. L'appendicite peut se présenter comme une occlusion fébrile avec une douleur chronique de la fosse iliaque droite associée à des troubles du transit. Ce tableau clinique peut suggérer une tumeur du côlon droit et impliquer une intervention chirurgicale (18) ;
- **appendicite de l'enfant et du nourrisson** : avant l'âge de 2 ans, l'appendicite est exceptionnelle. Elle peut néanmoins être gravissime en raison de son évolution rapide. L'âge habituel de survenue de l'appendicite dans l'enfance est entre 7 et 10 ans et le diagnostic sémiologique est identique à celui de l'adulte jeune (7) ;
- **les formes atypiques** : chez le patient diabétique ou le patient insuffisant rénal, les manifestations cliniques peuvent être abâtardies et le diagnostic retardé ;
- **les formes sous antibiotiques et antalgiques** : il n'est pas rare que des douleurs atypiques soient traitées de façon systématique par des antalgiques ou des antibiotiques. Il existe alors des formes abâtardies d'appendicites dont le diagnostic est difficile (7).

2.5 Diagnostic de l'appendicite

En raison du polymorphisme de son expression clinique et des localisations variables de l'appendice dans la cavité abdominale, l'appendicite reste un diagnostic qui peut être difficile.

En 1947, Henri Mondor décrivait dans son livre « les appendicites » et non « l'appendicite » (22).

2.5.1 Le diagnostic clinique et biologique

Le symptôme initial d'une appendicite est presque invariablement la douleur abdominale, de type viscéral, résultant des contractions de l'appendice ou de la distension de sa lumière. Elle est en général mal localisée dans la région péri-ombilicale ou péri-gastrique. C'est le seul signe constant (95% des cas) qui persiste lors des examens successifs même dans une forme très atypique (5,21).

Le diagnostic d'appendicite aiguë est d'abord clinique et repose sur des signes généraux et abdominaux notamment une fièvre supérieure à 38°C, une douleur provoquée à la palpation de la fosse iliaque droite et l'existence d'une défense pariétale (7,11,23).

La combinaison des résultats biologiques, notamment les signes de l'inflammation aiguë avec les caractéristiques cliniques, peut conduire à une meilleure précision diagnostique. Dans 80 à 85% des cas, il existe une hyperleucocytose à polynucléaires neutrophiles (> à $10 \cdot 10^9$ cellules /L). Cependant, celle-ci reste peu spécifique car si elle témoigne d'un processus inflammatoire, elle ne peut pas préjuger de la cause (11). En revanche, l'absence d'augmentation de la CRP (*C-Reactive Protéin*) et d'hyperleucocytose à polynucléaires neutrophiles témoigne d'une faible probabilité d'appendicite (11).

2.5.2 Les scores clinico-biologiques

Dans la littérature, plusieurs scores clinico-biologiques prédictifs d'appendicite aiguë ont été proposés. Ils sont établis statistiquement à partir de l'évolution d'un grand nombre de patients et ont pour objectif de prédire le diagnostic/pronostic du patient à partir de données cliniques et biologiques discriminantes.

Un tel système de score pourrait, s'il est performant, être la base d'un algorithme de prise en charge diagnostique qui permettrait la sélection des patients pour la chirurgie immédiate, l'observation ou les examens complémentaires d'imagerie. Il pourrait également être utilisé comme outil d'évaluation de l'évolution d'une appendicite sous traitement conservateur ou d'une appendicite résolutive. Il pourrait également permettre la caractérisation des patients inclus dans une étude clinique (24).

Néanmoins, la plupart des scores clinico-biologiques décrits dans la littérature sont élaborés soit à partir de séries rétrospectives, soit à partir de séries prospectives de patients déjà hospitalisés, leur qualité méthodologique en est par conséquent discutable (12,25).

Le score d'Alvarado est le plus décrit dans la littérature ; il est développé à partir d'une série rétrospective de patients hospitalisés pour suspicion d'appendicite. Il permet d'attribuer un score sur une échelle de 0 à 10 en combinant les éléments cliniques et biologiques.

En France, le score de Flamant a été retenu lors de l'élaboration des recommandations de l'ANDEM en 1996. Il permet à partir de trois critères clinico-biologiques simples de répondre à la question : « devant une douleur aiguë de la fosse iliaque droite, quels éléments permettent de surseoir à l'appendicectomie ? » (8,12).

2.5.3 L'imagerie

Dans un contexte clinique difficile, l'imagerie peut consolider une suspicion clinique, éliminer un diagnostic différentiel et rechercher d'éventuelles complications. L'imagerie contribue de manière significative au diagnostic de l'appendicite et à la prise en charge thérapeutique.

► L'ASP

La radiographie conventionnelle aux rayons X de l'abdomen sans préparation (ASP) a peu d'utilité dans ce contexte. Dans un rapport publié en janvier 2009, la HAS considère que l'ASP n'a plus d'indication dans les douleurs abdominales (26).

► L'échographie

L'échographie est utilisée depuis le début des années 80 pour le diagnostic de l'appendicite aiguë. Ces 30 dernières années, la technologie échographique a beaucoup évolué et la technique la plus communément utilisée pour les douleurs abdominales est l'échographie avec compression graduée.

L'échographie nécessite l'utilisation d'une sonde à haute fréquence ; les plus utilisées sont les sondes courbes (3,5-5,0 MHz) et les sondes linéaires (5,0-12,0 MHz). La fréquence utilisée dépend de la morphologie du patient.

Les critères échographiques de visualisation d'un appendice inflammatoire sont :

- un diamètre de plus de 6-8 mm avec une épaisseur pariétale de plus de 3mm ;
- un appendice non compressible ;
- une douleur provoquée par les manœuvres de compression.

Les avantages de l'échographie sont : le caractère non invasif, l'absence d'irradiation ou d'effets biologiques néfastes connus, la rapidité et les coûts relativement modestes. Elle permet de fournir, en temps réel, une information anatomique, fonctionnelle et dynamique avec un contrôle de l'image et une possibilité d'exploration des zones particulièrement sensibles.

De par son caractère non-irradiant, l'échographie est particulièrement utile chez l'enfant et la femme jeune. Son intérêt a également été montré chez cette dernière dans la recherche des diagnostics différentiels gynécologiques.

Les principaux inconvénients sont le caractère opérateur dépendant de l'acquisition et de l'interprétation de l'image et les limites inhérentes aux patients notamment la morphologie (obésité) et la distension aérique qui gêne la visualisation des structures digestives (27-29).

Plusieurs méta-analyses sont consacrées à l'évaluation de la performance diagnostique de l'échographie, elles sont analysées dans le cadre de ce travail.

► Le scanner

L'utilisation du scanner pour l'imagerie médicale est un phénomène relativement récent, le premier appareil disponible sur le marché date du début des années 70. Depuis, son utilisation dans les douleurs abdominales aiguës a pris beaucoup d'ampleur, notamment aux USA où entre 1996 et 2005 elle a augmenté de 141% (30).

La plupart des protocoles de scanner utilisent une acquisition hélicoïdale avec une épaisseur de coupe inférieure ou égale à 5 mm. Un scanner abdominal peut être réalisé avec ou sans produit de contraste iodé, par voie intraveineuse (IV) avec ou sans contraste oral/rectal (27). Il faut rappeler cependant que l'opacification digestive par voie haute n'est pas préconisée en France (31).

Le diagnostic d'appendicite est établi sur les critères suivants :

- un diamètre de plus de 6 mm avec une épaisseur pariétale de plus de 3mm ;
- des signes d'inflammation périappendiculaire identifiés sous la forme d'une densification de la graisse voire d'un abcès.

Les avantages du scanner sont la haute résolution des images et la précision diagnostique élevée en particulier pour les diagnostics différentiels (27-29).

Les inconvénients du scanner sont : le coût, le risque de réaction allergique au produit de contraste, la durée et son irradiation non négligeable (23). Pour un scanner abdominal, le niveau d'exposition moyen du patient est de III (cotation de 0 à IV) ce qui correspond en termes de dose efficace à une irradiation de l'ordre de 5 à 10 mSv. Cette dose correspond à une seule acquisition selon un protocole standard, dans le cas de plusieurs acquisitions successives sur le même volume, le niveau de dose est multiplié par le nombre d'acquisitions et l'examen passe donc fréquemment en classe IV (Source SFR) (32).

Les rayonnements du scanner sont responsables de 1,5 à 2% des cas de cancers aux Etats-Unis. Ce risque concerne en particulier, les patients souffrant de maladies chroniques ou récurrentes (33). Le risque additionnel d'un cancer radio-induit est quantifié de la manière suivante (34) : 0,5% pour 20 mSv, 0,6% pour 34 mSv et 1% pour 100 mSv.

Plusieurs méta-analyses sont consacrées à l'évaluation de la performance diagnostique du scanner ; elles seront analysées dans le cadre de ce travail.

► L'IRM

L'IRM est une modalité d'imagerie intéressante pour les patients chez qui les risques de rayonnement ou de néphrotoxicité potentielle des produits de contraste iodés présente un risque important, c'est en particulier le cas des femmes enceintes et de l'enfant.

L'échographie est l'examen de première intention en cas de suspicion d'appendicite chez la femme enceinte. Cependant, quand l'échographie n'est pas contributive ou a abouti à un examen incertain, certaines études récentes recommandent l'IRM (29,35).

Selon les critères de pertinence édictés par l'*American Collège of Radiology* (ACR), l'ACR considère, qu'en cas de douleur en fosse iliaque droite, de fièvre et de leucocytose chez la femme enceinte, l'IRM est « plus approprié » que le scanner. L'ACR considère également que l'IRM est une alternative intéressante à l'échographie pour les femmes enceintes et les enfants, dont l'exposition aux rayonnements ionisants relève d'une préoccupation majeure (36).

Dans le champ retenu pour ce travail, qui exclut la femme enceinte et l'enfant, il n'a pas été identifié de données probantes dans la littérature qui permettent d'évaluer la performance diagnostique de l'IRM chez les autres populations. De plus, compte-tenu du parc IRM Français, son utilisation dans cette indication ne semble pas réaliste.

2.5.4 L'exploration cœlioscopique

La cœlioscopie est l'une des modalités d'exploration morphologiques qui pourrait être proposée en cas de persistance de doute diagnostique devant une douleur abdominale aiguë. La place de cette modalité diagnostique dans la stratégie diagnostique sera évaluée dans ce travail.

3. Problématique et objectifs du travail

3.1 Définition de la problématique

Le diagnostic de l'appendicite est fréquent mais il reste difficile en raison du polymorphisme des symptômes, de l'absence de parallélisme anatomo-clinique et du caractère imprévisible de son évolution.

Les erreurs de diagnostic persistent encore aujourd'hui. Une étude américaine de 2001 rapporte que dans 15,5% des cas, l'appendice est normal lors de la chirurgie et ce taux peut atteindre jusqu'à 45% dans certains sous-groupes comme les femmes en âge de procréer, les personnes âgées et les enfants (37).

L'exérèse de l'appendice sain n'est pas sans conséquences ; l'exploration chirurgicale a sa morbidité propre et les complications occlusives du grêle sans relation avec le degré d'inflammation de l'appendice peuvent atteindre 3% (38).

Depuis 1999, le nombre total d'appendicectomies a beaucoup diminué en France, mais le nombre d'appendicectomies pour des appendicites compliquées a augmenté. Par ailleurs, une grande hétérogénéité interrégionale voire interdépartementale est constatée.

Au regard de ces éléments, il devrait encore être possible d'améliorer et d'optimiser l'indication de cet acte chirurgical.

L'enjeu est l'adoption d'une attitude diagnostique appropriée qui permette aussi bien de diminuer le risque des appendicites perforées que de réduire le nombre d'appendicectomies inappropriées.

3.2 Objectifs du travail

Les objectifs tels que validés par le collège de la HAS le 20 juillet 2012 (4) visent à produire les référentiels qui définissent la stratégie diagnostique permettant de poser l'indication d'une appendicectomie.

3.2.1 Questions incluses dans le champ de travail

Pour cela, il faudrait définir :

- la valeur diagnostique et la place dans l'aide à la décision chirurgicale :
 - des signes cliniques,
 - des signes biologiques,
 - des examens d'imagerie,
 - de la cœlioscopie diagnostique.
- l'intérêt potentiel des scores clinico-biologiques ;
- une stratégie diagnostique adaptée aux situations les plus courantes chez l'adulte (homme jeune, femme en âge de procréer, sujet âgé).

Le cas de la femme enceinte et de l'enfant de moins de 15 ans nécessitent une approche spécifique qui pourrait être envisagée dans un travail ultérieur.

Dans un second temps, la HAS pourrait définir un programme de mise en œuvre des référentiels qui associeraient selon les cas :

- l'élaboration d'un parcours type du patient à partir des processus de soins possibles ;
- la proposition pour chaque étape du parcours patient : des objectifs de qualité issus des référentiels et de données de la pratique (par exemple issues du PMSI), d'outils et méthodes de mesure et d'amélioration de la pertinence des interventions de santé (par exemple document d'information du patient, critères d'évaluation de la pratique professionnelle, dossier standardisé, aide-mémoire, arbre décisionnel, etc.).

3.2.2 Questions exclues du champ de travail

Dans ce travail, il ne s'agit pas de remettre en question la pertinence de réalisation d'une appendicectomie dans un contexte d'appendicite aiguë avérée, ni de comparer l'efficacité des deux techniques chirurgicales (laparotomie et cœlioscopie).

Il ne s'agit pas non plus de développer les outils de mesure de la pertinence de l'appendicectomie, travail qui sera réalisé en aval de ce travail.

Sont par conséquent exclus de ce travail :

- la comparaison de la laparotomie et de la cœlioscopie ;
- l'évaluation de l'efficacité de l'appendicectomie vs le traitement conservateur par antibiothérapie encore débattu aujourd'hui⁵ ;
- la définition des conditions de réalisation et des modalités d'hébergement (hospitalisation ou en ambulatoire) ;
- l'élaboration des critères de mesure d'impact de ces référentiels.

⁵ Dans la version du cadrage validée par le Collège de la HAS le 20 juillet 2011, cet item avait été exclu du champ d'évaluation. Cependant, en raison d'éléments récents de l'actualité, un chapitre consacré à la comparaison des deux traitements a été intégré à ce travail.

4. Méthode de travail

Ce travail s'inscrit dans un objectif d'amélioration de la qualité et de la sécurité des soins. Il s'agit d'établir des référentiels permettant aux différents acteurs de santé (professionnels, patients et usagers, décideurs) de disposer d'un outil qui aide :

- à la prise de décision dans le choix des soins ;
- à harmoniser les pratiques ;
- à réduire les traitements et actes inutiles ou à risque.

La méthode de réalisation retenue pour ce travail s'inspire de la méthode de « consensus formalisé », elle repose sur l'analyse critique et la synthèse des données bibliographiques et la participation de deux groupes multidisciplinaires de professionnels :

- un groupe d'orientation ;
- un groupe de cotation.

Le choix de cette méthode se justifie d'une part par le faible corpus de la littérature de bon niveau de preuves répondant spécifiquement aux questions et d'autre part, par la diversité des situations cliniques.

4.1 Analyse de la littérature

Compte-tenu de l'importance du champ de travail, de l'orientation de la demande dont l'objectif est de définir les éléments décisionnels permettant de poser le diagnostic d'une appendicite aiguë, la recherche sur *Medline* a été limitée dans un premier temps aux recommandations, méta-analyses et revues systématiques descriptives de langue anglaise et française, sans restriction sur la période de recherche.

Devant le faible corpus de la littérature identifiée avec ces critères, la recherche a été élargie aux études randomisées et aux études prospectives. La stratégie de recherche a été adaptée pour chaque item (signes cliniques, signes biologiques, imagerie...), notamment pour la période de recherche, en fonction de la disponibilité des données, de leur qualité méthodologique et de manière à compléter les données par rapport aux dates de fin de recherche rapportées dans les recommandations et méta-analyses identifiées lors de la première étape de recherche. De manière générale, cette deuxième recherche couvre la période 2004-2012.

4.1.1 Recherche documentaire

La recherche a été limitée aux publications en langue anglaise et française, elle a été faite en février 2011, sans limite de date inférieure. Une veille a été réalisée jusqu'à avril 2012.

Les sources suivantes ont été interrogées :

- pour la littérature internationale : les bases de données *Medline* ;
- la *Cochrane Library* ;
- les sites internet publiant des recommandations, des rapports d'évaluation technologique ou économique ;
- les sites Internet des sociétés savantes compétentes dans le domaine étudié.

Cette recherche a été complétée par les références citées dans les documents analysés.

La stratégie de recherche et la liste des sources interrogées sont détaillées dans l'Annexe 3.

► Documents identifiés

Le nombre de références identifiées dans les bases de données bibliographiques est de 759 et le nombre de références trouvées sur Internet est de 19 totalisant ainsi 778 références.

Première sélection

A la lecture et l'analyse des résumés des documents identifiés, une première sélection a été réalisée en fonction de l'objectif de l'étude, de son schéma et de la population étudiée.

Seules sont retenues, les méta-analyses, les revues systématiques et les études prospectives portant sur le diagnostic clinique, biologique ou macroscopique (imagerie et coelioscopie) de l'appendicite et s'adressant à une population adulte hospitalisée pour suspicion d'appendicite, douleur en fosse iliaque droite ou douleur abdominale aiguë.

Ont été par ailleurs exclues :

- les études qui portent sur les techniques chirurgicales ;
- les études rétrospectives, les revues générales, les éditoriaux, les études de cas et les lettres ;
- les études réalisées chez les enfants et la femme enceinte.

Au terme de cette première sélection, il ressort tous items confondus :

- 7 recommandations ;
- 29 méta-analyses ;
- 2 revues systématiques ;
- 46 autres types d'études.

Deuxième sélection : documents analysés dans ce rapport

Une lecture plus approfondie des articles a permis d'exclure :

- les études rétrospectives non identifiées à la lecture des résumés ;
- les études qui portent sur les patients opérés ;
- les études dont les résultats ne permettent pas d'évaluer les performances diagnostiques ;
- les études dont le sex-ratio est fortement déséquilibré (ne correspond pas aux ratios publiés dans la littérature soit un rapport femmes/hommes d'environ 1,3) ;
- Les études qui analysent la performance diagnostique du scanner avec une modalité de contraste oral (39). Cette modalité n'est pas conforme à la pratique française qui ne préconise pas l'opacification digestive par voie haute (31) ;
- de manière générale les études qui sont jugées comme méthodologiquement invalides (cf. : § 4.1.3).

En fonction de la nature des documents, des critères d'exclusion plus spécifiques ont été appliqués :

- les méta-analyses sont exclues si elles ne satisfont pas à au moins 8 critères sur 11 de la grille d'analyse AMSTAR et si elles présentent un biais de sélection géographique ;
- les études sur la performance diagnostique si la validité méthodologique est jugée discutable au regard des critères de la grille QUADAS-2 ;
- les études comparatives sont exclues si la comparaison des deux tests n'est pas réalisée chez le même patient.

Résultat final au terme de la sélection

Au terme de cette sélection et vis-à-vis de la question à traiter, les articles suivants sont retenus :

- signes cliniques : une méta-analyse, une revue systématique, une étude prospective et une recommandation ;
- signes biologiques : une méta-analyse, trois études prospectives et une recommandation ;
- scores clinico-biologiques : une revue systématique avec méta-analyse, une étude prospective et une étude contrôlée randomisée ;
- imagerie : quatre méta-analyses, trois recommandations ;
- coelioscopie diagnostique : une revue systématique, une conférence de consensus italienne, une recommandation des sociétés savantes françaises ;

- traitement conservateur : une revue systématique et une méta-analyse.

4.1.2 Analyse de la qualité méthodologique des documents

Pour chaque item⁶, les biais méthodologiques des articles sont analysés et décrits dans la partie argumentaire qui leur est consacrée.

Chaque article sélectionné est analysé selon les principes de lecture critique de la littérature en s'inspirant des grilles d'analyse ou de lecture suivantes :

► La grille AMSTAR

Il s'agit d'un questionnaire en onze éléments (cf. Annexe 4) qui permet d'évaluer la qualité méthodologique des revues systématiques et méta-analyses en évaluant les critères suivants :

- une conception *a priori* ;
- une sélection d'études et une extraction de données en double ;
- une recherche documentaire exhaustive ;
- l'utilisation de l'état de publication comme critère d'inclusion ;
- une liste d'études incluses ou exclues ;
- les caractéristiques des études incluses ;
- l'évaluation documentée de la qualité scientifique des études incluses ;
- l'utilisation appropriée de la qualité scientifique lorsqu'il s'agit de formuler des conclusions ;
- l'utilisation appropriée de méthodes pour combiner les résultats des études ;
- l'évaluation de la probabilité de biais de publication.

► La grille QUADAS

La lecture critique des articles sur la performance diagnostique s'est appuyée sur le principe de lecture de la grille QUADAS-2 (40). Cet outil est structuré selon quatre domaines, chacun disposant d'un ensemble de questions clés à rechercher et évaluer dans les articles pour juger des risques de biais et de l'applicabilité.

► La fiche de validité interne (Centre Cochrane Français)

La lecture critique des recommandations a été réalisée selon la fiche de validité interne publiée par le Centre Cochrane Français⁷. Cette fiche s'inspire du guide d'analyse de la littérature et gradation des recommandations de l'ANAES (41) et de la Grille d'évaluation de la qualité des recommandations pour la pratique clinique (Traduction française de la grille AGREE) (42).

► Critères de qualité d'une conférence de consensus (ANAES 1999)

L'évaluation de la qualité méthodologique d'une conférence de consensus repose sur les critères de qualité définis par l'ANAES en 1999 (43) pour l'élaboration de recommandations médicales ou professionnelles.

4.2 Position des professionnels de santé

Les experts ont été consultés selon une méthode de consensus formalisé simplifié⁸ conformément aux directives du « guide méthodologique » publiées en 2010 (44). Deux groupes de professionnels sont impliqués, leurs rôles de même que les modalités de constitution ont été validés par le collège de la HAS lors de l'examen de la note de cadrage consacrée à ce sujet (4).

⁶ Signes cliniques, signes biologiques, scores clinico-biologiques, imagerie, coelioscopie diagnostique

⁷ Centre Cochrane français : <http://tutoriel.fr.cochrane.org/fr>

⁸ Simplifications apportées : groupe de pilotage remplacé par un groupe d'orientation ; analyse critique de la littérature réalisée par la HAS ; et en vue de formaliser le degré d'accord des experts, relecture et cotation en un tour unique par un groupe de cotation.

4.2.1 Méthode de constitution des groupes de professionnels

La méthode de constitution des groupes d'orientation et de cotation est basée sur la représentativité des différentes professions et spécialités concernées et sur le respect, autant que possible, d'un équilibre du nombre de professionnels en fonction de leur mode d'exercice public ou privé, d'origine géographique ou d'écoles de pensée diverses.

Ces deux groupes sont composés par des professionnels de santé indiqués par les organismes professionnels suivants :

Spécialités	Organismes professionnels
Chirurgie viscérale et digestive	Fédération de chirurgie viscérale et digestive (FCVD) Société française de chirurgie digestive (SFCD)
Gastro-entérologie	Fédération des spécialistes des maladies de l'appareil digestif (FSMAD) Société nationale française de gastroentérologie (SNFGE)
Radiologie	Société française de radiologie (SFR)
	Société d'Imagerie Abdominale et Digestive (SIAD)
Médecine d'urgence	Collège Français de médecine d'urgence, (CFMU)
Biologie clinique	Société Française de Biologie Clinique (SFBC)
Anatomopathologie	Conseil national des pathologistes (CNP)
Médecine générale	Regroupement des sociétés scientifiques de médecine générale (RSSMG)
Gynécologie obstétrique	Collège national des gynécologues et obstétriciens français

Le Collège français d'anesthésie réanimation (CFAR) avait également été sollicités mais n'a pas indiqué de noms.

4.2.2 Groupe d'orientation

- Le groupe d'orientation a été sollicité pour :
 - contribuer à identifier, dans le cadre du champ d'évaluation défini par la HAS, les questions à traiter au regard de l'état de l'art et des difficultés rencontrées dans la pratique,
 - relire l'argumentaire scientifique rédigé par le chef de projet HAS après analyse critique et synthèse des données bibliographiques disponibles,
 - contribuer, sur la base de l'argumentaire scientifique et d'une discussion relative aux pratiques existantes, à formuler les propositions de référentiels et à élaborer l'arbre décisionnel à soumettre au groupe de cotation.

La composition nominative du groupe d'orientation est présentée à la fin de ce document.

4.2.3 Groupe de cotation

Le groupe de cotation a été sollicité pour donner un avis, après lecture de l'argumentaire scientifique, sur les propositions de référentiels et sur l'arbre décisionnel, proposés par le groupe d'orientation.

La composition nominative du groupe de cotation est présentée à la fin de ce document.

Les membres du groupe de cotation ont été interrogés par voie informatique via le portail de connexion Graal de la HAS. Chaque affirmation formulée, a été notée individuellement par un système de vote unique, chaque expert avait la possibilité d'émettre des commentaires dans un champ assorti à la proposition jugée. Les propositions faisant l'objet de consensus ont été utilisées

pour rédiger le texte final des référentiels, celles ne faisant pas l'objet de consensus ont été identifiées dans l'argumentaire.

Les règles de cotation fixées *a priori* et adressées à chaque participant sont les suivantes :

- le degré d'accord avec les conclusions énoncées dans le rapport est recueilli par la technique de cotation quantitative discontinue sur une échelle de 1 à 9 avec commentaires obligatoire pour les valeurs ≤ 3 :
 - la valeur 1 signifiait que l'affirmation présentée est « totalement inappropriée »,
 - la valeur 9 « totalement appropriée »,
 - la valeur 5 marque « l'indécision »,
 - Les valeurs 2 à 8 traduisent quant à elles l'ensemble des situations intermédiaires possibles.
- chaque professionnel⁹ doit impérativement compléter l'intégralité du questionnaire, en n'entourant pour chaque affirmation qu'un seul des chiffres présentés.

► Analyse des cotations par la HAS

Après gestion des valeurs manquantes et des valeurs extrêmes, le caractère de l'affirmation, ainsi que le degré d'accord sont définis en fonction de la dispersion et de la médiane des cotations de la manière suivante :

Proposition jugée		Valeur de la médiane	Distribution des réponses
Appropriée	Accord fort	≥ 7	Toutes les réponses comprises entre [7-9], sauf deux manquantes ou <7 .
	Accord relatif	≥ 7	Toutes les réponses comprises entre [5-9], sauf deux manquantes ou <5 , ou une manquante et une <5 .
Inappropriée	Accord fort	≤ 3	Toutes les réponses comprises entre [1-3], sauf deux manquantes ou >3 .
	Accord relatif	$\leq 3,5$	Toutes les réponses comprises entre [1-5], sauf deux manquantes ou >5 .
Incertaine	Indécision	Comprise dans [4-6,5]	Quelle que soit la répartition
	Absence de consensus	≥ 7	Au moins trois valeurs <5 ou manquantes.
		$\leq 3,5$	Au moins une valeur >5 ou manquantes

⁹ Un expert biologiste a souhaité, *a priori*, ne répondre qu'aux questions concernant son champ de compétence, sa voix n'a été prise en compte que pour les affirmations concernant la biologie.

4.2.4 Gestion des conflits d'intérêt

Les déclarations publiques d'intérêts (DPI) des membres des membres des groupes d'orientation et de cotation ont toutes été analysées selon le «Guide des déclarations d'intérêts et de gestion des conflits d'intérêts» de la HAS de mars 2010.

Le Bureau de la Commission d'évaluation économique et santé publique (CEESP) a jugé que les intérêts déclarés étaient compatibles avec la participation des différents experts qui composent les deux groupes. Aucun membre n'a déclaré d'intérêt majeur en relation avec le sujet de cette évaluation.

Le président de la société savante de médecine d'urgence considéré comme ayant un intérêt majeure de par sa position, n'a pas pu être retenu pour participer au groupe d'orientation ; il a cependant été auditionné par la HAS.

Toutes les DPI des experts participant à ce travail sont consultables sur le site de la HAS (www.has-sante.fr)¹⁰.

¹⁰ http://www.has-sante.fr/portail/upload/docs/application/pdf/guide_dpi.pdf

5. Analyse de la littérature

5.1 Questions traitées et limites méthodologiques

5.1.1 Questions traitées

Un symptôme, un signe clinique, le résultat d'un dosage biologique, d'un examen d'imagerie, sont autant d'éléments diagnostiques dont les résultats peuvent renseigner sur la probabilité plus ou moins grande d'une appendicite aiguë.

L'objectif de cette analyse est de définir la valeur diagnostique de chacun de ces éléments, celle de leur combinaison et de proposer une stratégie diagnostique adaptée aux situations cliniques.

Ce travail vise donc à documenter les points suivants :

- les signes cliniques les plus discriminants pour le diagnostic pour l'appendicite ;
- les paramètres biologiques les plus performants et leur valeur diagnostique pour l'appendicite ;
- la valeur diagnostique pour l'appendicite, de l'association des signes clinique et des examens biologiques ;
- la fiabilité des scores clinico-biologiques dans le diagnostic d'appendicite ;
- la performance diagnostique de l'échographie pour la détection d'une appendicite ;
- la performance diagnostique du scanner pour la détection d'une appendicite ;
- la place de la coelioscopie exploratrice dans la détection d'une appendicite ;
- l'identification des étapes permettant de poser le diagnostic d'appendicite.

5.1.2 Limites méthodologiques des études analysées

La lecture critique des articles analysés révèle d'importantes limites méthodologiques qui sont à l'origine d'une variabilité des paramètres d'exactitude diagnostique entre les différentes études.

Deux sources peuvent être à l'origine de cette variabilité : l'existence de variations réelles liées à l'hétérogénéité des caractéristiques démographiques générales ou des caractéristiques de la maladie dans la population et les biais qui résultent de la manière dont les études ont été conduites.

Plusieurs facteurs d'hétérogénéité sont identifiés dans les études :

- certaines études portent sur des patients se présentant avec une douleur abdominale aiguë, d'autres sur des patients avec suspicion d'appendicite (sans aucune précision du tableau clinique) et enfin d'autres études sont consacrées uniquement aux patients se présentant avec une douleur en fosse iliaque droite (FID). La variabilité de la prévalence d'appendicite dans ces trois types de populations constitue un facteur d'hétérogénéité susceptible d'affecter la performance diagnostique des différents tests ;
- l'appendicite étant une maladie évolutive, un examen réalisé sur un patient donné est une photographie à un temps t donné de l'état d'évolution de sa maladie. Il a été montré que la précision diagnostique chez un même patient dépendait du délai qui sépare le début de la douleur de la réalisation des examens (45,46). Or, au sein d'une population de patients recrutés dans une étude, différents niveaux d'évolution et de gravité sont représentés alors même que la distribution de la gravité de la maladie dans l'échantillon influence la sensibilité du test ;
- il existe un grand polymorphisme des profils clinico-biologiques, tous les signes ne se retrouvent pas systématiquement chez tous les patients. Plusieurs associations sont possibles et il est très difficile de décrire un profil « type » d'appendicite (47) ;

- l'évaluation de la performance diagnostique des signes cliniques de même que celle des examens d'imagerie, pose le problème de la subjectivité de l'examen qui reste examinateur dépendant et donc difficilement standardisable. Cet aspect engendre indéniablement une hétérogénéité entre les études et conditionne la valeur diagnostique des tests ;
- pour le diagnostic d'imagerie, les données sont souvent simplifiées sous forme de données binaires (oui/non), analysées sur la base des données déclaratives du radiologue. Ces données sont insuffisantes pour permettre l'analyse de différents effets de seuils ou des paramètres recueillis sur la positivité des résultats notamment pour les cas indéterminés ou équivoques. De plus, quand ces seuils ou paramètres de positivité sont renseignés ils peuvent être très variables au sein d'une même étude, à titre d'exemple, un diamètre de paroi de l'appendice de 6 mm est un critère de positivité à l'échographie mais d'autres critères sont également utilisés notamment l'aspect apéristaltique ou l'œdème péri-appendiculaire, etc. (48).

Plusieurs biais sont également identifiés dans les études :

- le principal réside dans le choix du test de référence, critère essentiel pour témoigner de la présence ou de l'absence de la maladie. Or, dans le cas de l'appendicite, le résultat de l'analyse histopathologique sur pièce d'exérèse reste un outil imparfait pour plusieurs raisons :
 - ▶ il ne peut être appliqué qu'aux patients opérés, ce qui représente un biais de vérification différentiel,
 - ▶ les critères histopathologiques de l'appendicite sont controversés dans la littérature, une infiltration inflammatoire de la muqueuse peut être retrouvée dans près de 35% des cas, elle est acceptée par certains comme signe précoce de la maladie alors que d'autres estiment qu'une inflammation transmurale est requise pour le diagnostic d'une appendicite (49),
 - ▶ il existe des appendicites spontanément résolutive (8%) décrites dans des études de suivi par imagerie échographique et tomodensitométrie (50).
- dans la majorité des études, les résultats du suivi des patients pour qui le retour à domicile est décidé (patients non opérés) ne sont pas documentés ;
- il existe souvent un biais d'absence d'insu pour l'évaluation de la performance diagnostique de l'imagerie car l'examen anatomopathologique est rarement réalisé en insu des résultats clinique et de l'imagerie. Cette absence d'insu peut favoriser la concordance entre le test de référence et le résultat de l'examen d'imagerie.

Tous ces facteurs sont autant d'éléments susceptibles d'impacter la valeur diagnostique des éléments cliniques, biologiques et des examens d'imagerie.

5.2 Les signes cliniques et biologiques

Dans ce travail, les données recueillies dans les études pour décrire les performances diagnostiques d'une variable sont les aires sous la courbe ROC et les rapports de vraisemblance. Ces mesures sont mieux adaptées que la sensibilité et la spécificité pour décrire et comparer les performances des variables diagnostiques de faible pouvoir discriminant. Le rapport de vraisemblance donne une mesure indirecte en renforçant la probabilité de la maladie quand il est supérieur à 1 et en diminuant la probabilité quand il est inférieur à 1. Plus le rapport de vraisemblance positif est élevé et plus le rapport de vraisemblance négatif est faible, plus le gain diagnostique du test est important (cf. Annexe 1).

5.2.1 Valeur diagnostique des signes et symptômes cliniques

Il existe très peu d'études qui évaluent la performance diagnostique des signes et symptômes clinique de l'appendicite, la recherche de la littérature a permis d'identifier pour ce paramètre, une méta-analyse (51), une revue systématique (52), une recommandation internationale de l'ACEP (53) et une seule étude prospective (47).

Toutes les études présentent de nombreuses limites méthodologiques qui les rendraient inéligibles au regard des grilles de lecture utilisés. Néanmoins étant donné le manque de littérature sur ce

paramètre, ces articles sont retenus et les résultats considérés comme exploratoires sont présentés à titre descriptif.

► Analyse des articles

Méta-analyse d'ANDERSON 2004

L'objectif de la méta-analyse d'Anderson (51), était l'analyse de la valeur diagnostique des éléments relatifs à l'histoire de la maladie, aux symptômes et signes cliniques et aux variables biologiques chez les patients admis à l'hôpital pour suspicion d'appendicite. L'analyse avait regroupé les résultats de 24 études prospectives consécutives totalisant 5833 patients.

Le principal critère de jugement était l'examen histopathologique pour les patients opérés. Pour les patients non opérés le critère de jugement était l'absence de survenue d'appendicite pendant le suivi. Seules cinq études sur 24 ont rapporté les résultats du suivi clinique pour ces patients (cf. tableau 1 annexe 9).

La prévalence d'appendicite dans les études variait entre 26,7 et 60,6 % avec une valeur médiane autour de 41%. Le taux d'appendicites perforées (rapporté dans 13 études) variait entre 3,7 et 28,6% avec un taux médian de 17,4% et le taux d'appendicites négatives (rapporté dans 19 études) variait de 5,2 à 42,2% avec un taux médian autour de 20%.

Les seuls symptômes et signes cliniques qui semblent présenter un pouvoir discriminant significatif ($p < 0,001$), qui reste néanmoins modéré (entre 0,68 et 0,70 sur la courbe ROC) (cf. tableau 1), sont :

- la douleur migrante ;
- la défense abdominale ;
- la douleur à la décompression en fosse iliaque.

Les RVP pour ces 3 signes respectivement de 2,06, 2,48 et 1,99 ($p < 0,001$), montrent une faible performance diagnostique. Les RVN compris entre 0,39 et 0,52 ne permettent pas d'exclure une appendicite en absence de ces signes.

De nombreuses limites méthodologiques (cf. annexe 9 tableau 1 et grille AMSTAR), ne permettent pas d'attribuer à cette étude un niveau d'évidence élevé.

Revue systématique de Wagner 1996

Il s'agit d'une revue systématique avec méta-analyse dont l'objectif était d'évaluer la valeur discriminante des éléments de l'histoire de la maladie et de l'examen clinique pour le diagnostic de l'appendicite chez une population adulte (52). Le test de référence était l'examen histopathologique (cf. tableau 2 annexe 9).

Au total, 11 études (5275 patients) ont été analysées, cinq incluaient une population avec une douleur abdominale aiguë évoluant depuis moins d'une semaine (2389 patients), les six autres études incluaient des patients avec suspicion d'appendicite mais sans précision des critères d'inclusion (2886 patients).

L'âge des patients est présenté sous forme d'intervalles dans les études, avec une borne inférieure variant de 4 à 15 ans et une borne supérieure variant de 50 à 90 ans. Le pourcentage moyen de femmes dans la méta-analyse est de 41%, il varie de entre 39 à 58% selon les études avec une étude de 1000 patients dédiée exclusivement à une population masculine (0% de femmes). Le taux de prévalence d'appendicite n'est pas renseigné ;

Les auteurs ont rapporté que les données recueillies dans les études étaient « similaires » et que les rapports de vraisemblance ont été calculés selon un modèle à effets fixes avec un intervalle de confiance à 95%. Quand une hétérogénéité a été détectée, les résultats ont été rapportés sous forme d'intervalle.

Dans cette étude, trois signes cliniques avec un RVP compris entre 3,18 et 8 (tableau 2) suggèrent quand l'un d'eux est présent un diagnostic d'appendicite :

- une douleur en FID ;
- une douleur migrante ;
- une défense.

En revanche, aucun signe considéré individuellement ne présentait un RVN suffisamment important (<0,2) pour permettre quand il est absent d'exclure le diagnostic d'une appendicite.

L'étude conclut qu'il est difficile d'affirmer ou d'exclure le diagnostic d'une appendicite sur simple examen clinique et qu'il est prudent de réaliser un suivi étroit des patient souffrant de douleurs abdominales tant que d'autres examens complémentaires n'ont pas été réalisés.

De nombreuses limites méthodologiques (cf. annexe 9 tableau 2 et grille AMSTAR), ne permettent pas d'attribuer à cette étude un niveau d'évidence élevé.

Etude prospective

Une seule étude prospective consécutive multicentrique de 942 patients dont l'objectif était d'évaluer la précision diagnostique des caractéristiques cliniques et des variables biologiques pour le diagnostic de l'appendicite aiguë a été identifiée (47). Il s'agit d'une sous étude pré-planifiée dont les données ont été recueillies dans le cadre d'un protocole prospectif multicentrique qui évaluait la valeur ajoutée de l'imagerie (radiographie plus échographie plus scanner) après évaluation clinique et biologique chez les patients se présentant aux urgences avec une douleur abdominale aiguë.

Le diagnostic final était établi par un groupe d'experts sur la base du recueil, chez tous les patients de l'étude, des éléments ci-dessous :

- les signes et symptômes cliniques¹¹ ;
- la numération sanguine ;
- le dosage de la CRP ;
- l'examen anatomopathologique (pour les patients opérés uniquement) ;
- le rapport d'imagerie ;
- les éléments de suivi clinique sur 6 mois.

L'âge moyen des patients dans l'étude était de $47 \pm 17,3$ ans dont 55 % de femmes, ils présentaient tous une douleur abdominale aiguë évoluant depuis au moins 2 h et moins de 5 jours. Les données analysées, correspondent à un faisceau de variables diagnostiques¹² identifiées dans la littérature notamment dans la méta-analyse d'Anderson et dans la revue systématique de Wagner 1996 décrites plus haut.

Quelle que soit la variable, la valeur de l'aire sous la courbe ROC est comprise entre de 0,50 à 0,65 ce qui témoigne d'une faible capacité discriminante individuelle des variables diagnostiques.

Sur le plan méthodologique, cette étude observationnelle souffre d'un biais de sélection, il s'agit d'une sous étude d'un protocole construit pour évaluer la performance diagnostique de la radiographie de l'échographie et du scanner et les patients qui ne nécessitaient pas un recours à l'imagerie ne sont pas inclus dans l'étude.

¹¹ L'histoire de la douleur en FID, la douleur migrante en FID, la douleur lors du mouvement, l'anorexie, la durée des symptômes, la nausée, la douleur progressive, l'histoire de la fièvre, les vomissements, la diarrhée, la sensibilité en FID, la douleur à la décompression, la sensibilité abdominale, la sensibilité rectale, la température.

¹² L'histoire de la douleur en FID, la douleur migrante en FID, la douleur lors du mouvement, l'anorexie, la durée des symptômes, la nausée, la douleur progressive, l'histoire de la fièvre, les vomissements, la diarrhée, la sensibilité en FID, la douleur à la décompression, la sensibilité abdominale, la sensibilité rectale, la température.

► Les recommandations

En 2010, l'American College of Emergency Physicians l'ACEP (53) a édité des recommandations dont l'objectif principal était de répondre à des questions relatives à l'évaluation et à la gestion des patients admis au département des urgences pour suspicion d'appendicite.

Au travers de ce travail l'ACEP visait entre autres à répondre à la question suivante :

« Les résultats cliniques peuvent-ils être utilisés pour stratifier le risque des patients d'avoir une appendicite et guider la prise de décision ? »

La méthode d'élaboration de ces recommandations reposait sur l'analyse de la littérature existante et sur un consensus des urgentistes quand la littérature n'était pas disponible. La recherche de la littérature a été effectuée sur la base MEDLINE et la base de données Cochrane, elle couvrait la période Janvier 2000 à Mars 2007. Les données cliniques et les recommandations ont été discutées et classés par niveau de preuves (de 1 à 3 cf. annexe 9) en tenant compte de la qualité méthodologique des articles sur lesquels reposaient les conclusions.

La recommandation s'adresse à une population de patients adultes se présentant avec une douleur abdominale aiguë non traumatique avec suspicion d'appendicite.

En ce qui concerne la question relative aux éléments de diagnostic clinique, la recommandation de l'ACEP a reposé principalement sur les résultats de la méta-analyse d'Anderson (51) et sur la revue systématique de Wagner et al. (52) présentées plus haut et sur deux études de faible niveau de preuves (niveau 3). Les principaux signes et symptômes décrits sont notamment ceux identifiés précédemment (51,52) il s'agit principalement de :

- la douleur en fosse iliaque droite (FID) ;
- la douleur migrante ;
- la sensibilité en FID ;
- la défense abdominale.

L'ACEP recommande :

« Chez les patients avec suspicion d'appendicite aiguë, « utilisez » l'examen clinique (signes et symptômes) pour stratifier le risque et orienter la décision concernant les examens complémentaires (pas d'examens, examens biologiques, et/ou imagerie) et celle concernant la prise en charge ultérieure du patient (sortie, observation et/ou consultation chirurgicale)» niveau B

Cette formulation est telle qu'elle ne définit pas clairement les éléments cliniques qui permettent de caractériser une « appendicite ». Elle confirme tout au plus le rôle de l'histoire de la maladie et de l'examen clinique dans la stratification du risque d'appendicite. Ceci est justifié par le manque de robustesse des éléments issus des principales études sur lesquelles reposent les affirmations.

Tableau 1. Signes et symptômes cliniques, capacité discriminante (AUC) et performance diagnostique (RVP et RVN)

Auteur, Année, référence		Symptômes et signes cliniques	Nombre de patients (n études)	ROC	p	RVP [IC 95%]	p	RVN [IC 95%]	p
Anderson, 2004 (51)	Histoire de la maladie	- Douleur migrante	2459 [7]	0,68 [0,63-0,74]	p<0,001	2,06 [1,63-2,60]	p<0,001	0,52 [0,40-0,69]	p<0,001
		- Défense	2004 [6]	0,68 [0,60-0,76]	p<0,001	2,48 [1,60-3,84]	p<0,001	0,57 [0,48-0,68]	P=0,015
	Examen clinique	- Douleur à la décompression	3439 [9]	0,70 [0,65-0,75]	p<0,001	1,99 [1,61-2,45]	p<0,001	0,39 [0,32-0,48]	p=0,004
Wagner et al., 1996 (52)	Histoire de la maladie	- Douleur en FID	3979 [8]	NR		7,31-8,46*	NR	0-0,28*	P>0,05
		- Douleur migrante	1354 [5]	NR		3,18 [2,41-4,21]		0,50 [0,42-0,59]	P=0,05
	Examen clinique	- Défense	3555 [7]	NR			NR		P=0,05
		- Signe du Psoas	450 [3]	NR		3,76 [2,96-4,78]	NR	0,82 [0,79-0,85]	P=0,05
		- Fièvre	1264 [3]	NR		2,38 [1,21-4,67]	NR	0,90 [0,83-0,98]	P=0,05
							1,94 [1,63-2,32]		0,58 [0,51-0,67]

RVP : rapport de vraisemblance positive, RVN : rapport de vraisemblance négative, IC : intervalle de confiance, FID : Fosse iliaque droite, * en raison d'une grande hétérogénéité entre les études, les valeurs des rapports de vraisemblance sont présentées sous forme d'intervalle ; NR : non renseigné, NA : non applicable

5.2.2 Valeur diagnostique de la fièvre

La valeur diagnostique de la fièvre a été très peu étudiée, l'étude de Cardall et al. (54) rapporte pour une fièvre > 99°F (soit 37,2°C) une AUC ROC de 0,59 IC 95% [0,52-0,66] et des rapports de vraisemblance RVP de 1,3 IC 95% [0,97-1,72] et RVN de 0,82 IC 95% [0,65-1,01], valeurs qui témoignent d'une faible valeur diagnostique. Les caractéristiques et limites spécifiques de cette étude sont décrites dans le tableau d'évidence 4 en annexe 9.

Dans la revue systématique de Wagner et al. (52), la valeur diagnostique de la fièvre n'a été rapportée que dans 3 études sans précision du degré de température. Les rapports de vraisemblance RVP 1,94 IC 95% [1,63-2,32] et RVN 0,58 IC 95% [0,51-0,67] ne permettent pas d'attribuer une valeur diagnostique élevée à ce signe considéré individuellement.

5.2.3 Valeur diagnostique du sexe

Dans la méta-analyse d'Anderson (51) le RVP pour le sexe masculin est de 1,62 [1,49-1,76] IC 95%.

Dans l'étude de Laméris et al. (47), l'un des facteurs qui ressort comme ayant la plus forte valeur prédictive est le sexe masculin avec un RVP de 2 et une RVN de 0,6 (cf. annexe 9 tableau 3). Dans cette étude, les patients avec suspicion d'appendicite ont été classés par les auteurs en 23 profils différents, en fonction de l'âge, du sexe et des caractéristiques cliniques et biologiques. Pour la forme d'appendicite dite « classique » qui combine la migration de la douleur en FID, la sensibilité en FID, et la défense, la précision diagnostique est de 100% chez les hommes et seulement 46 % chez les femmes (cf. tableau 8).

► Conclusion

Au regard du faible niveau de preuves des articles et à titre uniquement descriptif, les éléments de diagnostic clinique qui ressortent de l'analyse de la littérature, sont :

- la douleur en fosse iliaque droite ;
- la douleur migrante ;
- la défense abdominale ;
- douleur à la décompression en fosse iliaque.

Tous ces signes et symptômes cliniques montrent individuellement un faible pouvoir discriminant et une performance diagnostique relativement faible qui ne permet pas d'affirmer ou d'exclure un diagnostic d'appendicite.

5.2.4 Valeur diagnostique des marqueurs biologiques de l'inflammation

Les marqueurs de l'inflammation ne peuvent être considérés comme spécifiques de l'appendicite. Néanmoins, c'est un outil diagnostique qui peut contribuer à augmenter la probabilité d'une appendicite. Dans ce travail et au travers de l'analyse de la littérature, les éléments biologiques qui sont identifiés et analysés sont principalement :

- le nombre de leucocytes ;
- le nombre de granulocytes ;
- le taux de polynucléaires neutrophiles (PNN) ;
- la valeur de la CRP.

► Analyse des articles

Cinq articles identifiés ont évalué la précision diagnostique des signes biologiques de l'inflammation, il s'agit de la méta-analyse décrite plus haut (51), de deux études prospectives consécutives monocentriques (46,54), d'une étude prospective multicentrique (47) et de la recom-

mandation internationale de l'ACEP (53). Hormis la méta-analyse, aucune étude n'évalue simultanément les 4 variables.

Pour les études prospectives et de manière générale, une des principales limites méthodologiques relevées est le biais de vérification différentiel (décrite en § 5.1.2.). Les limites spécifiques à chaque étude sont décrites dans les tableaux d'évidence en annexe 9. Au regard de la grille d'évaluation Quadas-2, toutes ces études présentent une faible qualité méthodologique qui les classe d'emblée en faible niveau de preuves.

► Caractéristiques des patients

Les caractéristiques des patients inclus dans les études sont très variables (cf. tableau 2). Dans les études d'Anderson (51) et de Cardall et al. (54), il s'agit de patients admis pour suspicion d'appendicite. L'étude de Laméris et al. (47) s'adressait à des patients souffrant d'une douleur abdominale aiguë non spécifique alors que l'étude de Ortega-Deballon et al. (46), s'adressait à des patients qui avaient une douleur localisée en FID.

Cette hétérogénéité des populations transparaît dans la variabilité des prévalences d'appendicites entre les études : autour de 65,7% dans l'étude d'Ortega-Deballon et al. un taux médian de 41 % IC 95 % [26,7-60,6%] dans la méta-analyse et une prévalence respectivement de 30% et 31 % dans les études de Laméris et al. (47) et de Cardall et al. (54),

L'âge moyen des patients était compris entre 30 et 47 ans avec un sex-ratio (femmes/hommes) comparable pour les 3 études (de 1,22 à 1,25), ces deux données ne sont pas renseignées dans la méta-analyse d'Anderson.

Tableau 2. Caractéristiques des patients dans les études

Auteur, année, référence,	Présentation clinique	Age moyen	Sex-ratio (femme/hommes)	Prévalence d'appendicite
Laméris <i>et al.</i> , 2009 (47)	DAA	47 ±17,3 [19-94]	1,22	30%
Ortega-Deballon <i>et al.</i> , 2008 (46)	Douleur en FID	33 [31,7-34,3]	1,23	65,7%
Cardall <i>et al.</i> , 2004 (54)	Suspicion d'appendicite	30,8 [7-75]	1,25	31%
Anderson, 2004 (51) Méta-analyse	Suspicion d'appendicite	NR	NR	41%

DAA : douleur abdominale aiguë ; FID : fosse iliaque droite ; NR : non renseigné

► Performance diagnostique des marqueurs biologiques de l'inflammation

Nombre de leucocytes

Dans le tableau 3 sont présentées les différentes valeurs des aires sous la courbes ROC (AUC ROC) obtenues à différentes concentration de leucocytes.

Pour une concentration standard ($\geq 10 \times 10^9$ cellules/L), la valeur de l'AUC ROC varie de 0,62 à 0,77 selon les études, ce qui témoigne d'une capacité discriminante relativement faible. Les valeurs des rapports de vraisemblance montrent également une performance diagnostique faible avec des RVP compris entre 1,4 et 2,47 et des RVN compris entre 0,26 et 0,4.

Dans les cas d'appendicites perforées (tableau 3), les leucocytes ($\geq 10 \times 10^9$ cellules/L) ont une capacité discriminante plus élevée, avec une AUC ROC de 0,85 et des valeurs de RVP et RVN respectivement de 4,2 et 0,2 témoignant d'une bonne performance diagnostique, en particulier

pour exclure une appendicite perforée quand le nombre de leucocytes est inférieur à 10×10^9 cellules/L.

Nombre de granulocytes

Le nombre de granulocytes en tant qu'élément prédictif d'appendicite est très peu décrit dans les études. Dans sa méta-analyse, Anderson (51) (tableau 4) rapporte pour une concentration supérieure ou égale à 13×10^9 cellules/L un RVP de 7,09, IC 95% [4,06-12,37] ce qui peut suggérer une forte probabilité d'appendicite, mais parallèlement la valeur du RVN est faible 0,74, IC 95% [0,68-0,81] et ne permet pas d'exclure le diagnostic d'appendicite en cas de normalité.

Pour les appendicites perforées, le nombre de granulocytes a une capacité discriminante relativement élevée avec une AUC ROC de 0,86 IC 95% [0,81-0,90]. La performance diagnostique pour une concentration supérieure ou égale à 7×10^9 cellules/L est élevée avec un RVP de 2,89 IC 95% [2,41-3,46] et un RVN de 0,14 IC 95% [0,08-0,26] ce qui signifie qu'un patient avec un nombre de granulocytes inférieur à 7×10^9 cellules/L a très peu de chance d'avoir une appendicite perforée (51).

Taux de polynucléaires neutrophiles (PNN)

Dans la méta-analyse d'Anderson, pour un taux supérieur à 75% de PNN, l'AUC ROC est de 0,77 IC 95% [0,70-0,84], le RVP est de 2,44 IC 95% [1,60-3,74] et le RVN de 0,24 IC 95% [0,11-0,50] (cf. tableau 5) (51).

Dans l'étude d'Ortega-Deballon et al. (46) pour un taux de PNN de 79,9%, l'AUC ROC est de 0,68 ($p < 0,001$), la valeur des rapports de vraisemblance (calculées à partir des valeurs moyennes de Se et Sp) sont de 1,52 pour le RVP et de 0,38 pour le RVN, ce qui montre une capacité discriminante et une performance diagnostique faibles pour ce paramètre.

La protéine C réactive (CRP)

Dans l'étude de Laméris et al. (47) (tableau 6), la capacité discriminante de la CRP pour une concentration supérieure ou égale à 10 mg/L est très faible AUC ROC de 0,55 IC 95% [0,49-0,60], les rapports de vraisemblance positifs et négatifs montrent également une faible performance diagnostique avec des valeurs respectivement de 1,1 et 0,7.

Pour la même concentration, la méta-analyse d'Anderson montre une valeur médiane d'AUC ROC de 0,75 IC 95%, [0,66-0,85], un RVP de 1,97 IC 95% [1,58-2,45] et un RVN de 0,32 IC 95% [0,20-0,51]. Pour les appendicites perforées la capacité discriminante est plus élevée, avec une AUC ROC de 0,87 IC 95% [0,74-1,01], un RVP de 4,24 IC 95% [1,16-15,53] et un RVN de 0,11 IC 95% [0,05-0,25].

Dans l'étude d'Ortega-Deballon et al. (46), la capacité diagnostique de la CRP semble plus élevée que dans les deux autres études, elle est également plus élevée que les autres signes inflammatoires individuels que sont les leucocytes et le taux de PNN (cf. tableaux 3 et tableau 5) avec une AUC ROC de 0,84 ($p < 0,0005$) et des RVP et RVN respectivement de 3,53 et 0,122 (calculés à partir des valeurs de sensibilité et spécificité moyennes présentées dans l'étude). Cette différence avec les valeurs rapportées par les deux autres études analysées et en règle générale par rapport aux valeurs constatées dans la littérature s'explique par la population très spécifique de cette étude (patients avec douleur abdominale aiguë en FID) qui présente une prévalence d'appendicite élevée (65,7%). Cette étude (46) montre également une corrélation entre la sévérité de l'appendicite et l'augmentation de la CRP, elle varie comme suit : pas d'appendicite 12.2 mg/L [3.9-20.5] $p < 0,0005$, appendicite phlegmoneuse 27 mg/L [18.9-35.1] $p < 0,0005$, appendicite gangréneuse 60.6 mg/L [41.1-80] $p < 0,0005$ et appendicite perforée 174.8 mg/L [107.8-241.8] $p < 0,0005$.

Au regard des données analysées, il apparaît une variabilité entre les études en ce qui concerne la capacité discriminante de la CRP, qui reste par ailleurs faible.

Influence de délai d'évolution de l'appendicite sur la performance diagnostique des signes inflammatoires

Dans l'étude d'Ortega-Deballon et al. (46), il a été recherché l'influence du délai, séparant l'apparition de la douleur de la réalisation des examens, sur la capacité discriminante des marqueurs biologiques de l'inflammation. Pour cela, les patients ont été stratifiés en fonction de la durée d'évolution : plus de 12 heures et moins de 12 heures d'évolution depuis l'apparition de la douleur.

La CRP et les leucocytes, montrent une meilleure précision diagnostique chez les patients ayant plus de 12 heures d'évolution :

- pour la CRP l'AUC ROC est de 0,896 vs 0,779 pour un délai inférieur à 12 h ($p < 0,0005$) ;
- pour les leucocytes l'AUC ROC est de 0,795 vs 0,609 pour un délai inférieur à 12 h ($p < 0,0005$).

La combinaison de la CRP et des leucocytes montre également une plus grande précision chez les patients avec plus de 12 heures d'évolution (0,913 vs 0,845, $p < 0,0005$).

Tableau 3. Nombre de leucocytes, capacité discriminante (AUC) et performance diagnostique (RVP et RVN)

Auteur, année, référence,	Présentation clinique (N)	Nombre de leucocytes	AUC ROC [IC95%]	RVP [IC95%]	RVN [IC95%]
Laméris <i>et al.</i> , 2009 (47)	DAA (942)	$>10 \times 10^9 / L$	0,62 [NR]	1,4 [NR]	0,4 [NR]
Ortega-Deballon <i>et al.</i> , 2008 (46)	Douleur en FID (134)	$14,8 \times 10^9 / L$	0,75 [NR]	1,5* [NR]	0,325* [NR]
Cardall <i>et al.</i> , 2004 (54)	Suspicion d'appendicite (293)	$>10 \times 10^9 / L$	0,72 [0,65-0,79]	1,59 [1,31-1,93]	0,46 [0,31-0,67]
Anderson, 2004 (51) Méta-analyse	Suspicion d'appendicite (3382)	$\geq 10 \times 10^9 / L$	0,77 [0,75-0,78] 0,85 [0,81-0,89]	2,47 [2,06-2,95]	0,26 [0,18-0,36] 0,48 [0,41-0,55] 0,69 [0,55-0,86] 0,81 [0,69-0,95]
		$\geq 12 \times 10^9 / L$		2,75 [1,99-3,80]	
		$\geq 14 \times 10^9 / L$		2,96 [2,48-3,53]	
		$\geq 15 \times 10^9 / L$		3,47 [1,55-7,77]	
	Appendicites perforées	$\geq 10 \times 10^9 / L$	0,20 [0,10-0,41] 0,6[0,56-0,78]	4,20 [2,11-8,35]	
		$\geq 15 \times 10^9 / L$		7,20 [4,31-12]	

* calculé à partir des valeurs de Se et Sp, **NR** : non renseigné, **RVP** : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients, **FID** : fosse iliaque droite, **DAA** : douleur abdominale aiguë

Tableau 4. Nombre de granulocytes, capacité discriminante (AUC) et performance diagnostique (RVP et RVN)

Auteur, année, référence,	Présentation clinique (N)	Nombre de granulocytes	AUC ROC [IC95%]	RVP [IC95%]	RVN [IC95%]
Anderson, 2004 (51)	Suspicion d'appendicite (3382)	$\geq 7 \times 10^9 / L$	0,78 (0,75-0,82)	1,64 [0,87-3,09]	0,31 [0,23-0,40] 0,45 [0,37-0,54] 0,60 [0,53-0,69] 0,74 [0,68-0,81]
		$\geq 9 \times 10^9 / L$		2,66 [1,39-5,09]	
		$\geq 11 \times 10^9 / L$		4,36 [2,83-6,73]	
		$\geq 13 \times 10^9 / L$		7,09 [4,06-12,37]	
	Appendicites perforées	$\geq 7 \times 10^9 / L$	0,86 [0,81-0,90]	2,89 [2,41-3,46]	
		$\geq 9 \times 10^9 / L$		4,16 [3,15-5,51]	

RVP : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients,

Tableau 5. Taux PNN, capacité discriminante (AUC) et performance diagnostique (RVP et RVN)

Auteur, année, référence,	Présentation clinique (N)	Taux de PNN	AUC ROC [IC95%]	RVP [IC95%]	RVN [IC95%]
Ortega-Deballon <i>et al.</i> , 2008 (46)	Douleur en FID (134)	79,9 %	0,68 [NR]	1,52*	0,376*
Anderson, 2004 (51)	Suspicion d'appendicite (3382)	>75% >85%	NR	2,44 [1,60-3,74] 3,82 [2,86-5,08]	0,24 [0,11-0,50] 0,58 [0,51-0,66]

* calculé à partir des valeurs moyennes de Se et Sp, IC non disponibles, **PNN** : polynucléaire neutrophiles, **RVP** : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients, **FID** : fosse iliaque droite

Tableau 6. CRP, capacité discriminante (AUC) et performance diagnostique (RVP et RVN)

Auteur, année, référence,	Patients (N)	CRP mg/L	AUC ROC [IC95%]	RVP [IC95%]	RVN [IC95%]
Laméris <i>et al.</i> , 2009 (47)	Douleur abdominale 942	>10	0,55 [0,49-0,60]	1,1 [NR]	0,7 [NR]
Ortega-Deballon <i>et al.</i> , 2008 (46)	Douleur en FID 134	67,7	0,846 p<0,0005	3,53*	0,122*
Anderson, 2004 (51)	Suspicion d'appendicite 3382	>10 >20	0,75 [0,66-0,85]	1,97 [1,58-2,45] 2,39 [1,67-3,41]	0,32 [0,20-0,51] 0,47 [0,28-0,81]
	Appendicites perforées	>10	0,87 [0,74-1,01]	4,24 [1,16-15,53]	0,11 [0,05-0,25]

* calculé à partir des valeurs moyennes de Se et Sp, **PNN** : polynucléaire neutrophiles, **RVP** : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients, **FID** : fosse iliaque droite, **NR** non renseigné

► Les recommandations

La seule recommandation identifiée pour cet item, est celle de l'ACEP (53), qui pour répondre à la première question posée précédemment (§ 5.2.1) : « Les résultats cliniques peuvent-ils être utilisés pour stratifier le risque des patients d'avoir une appendicite et guider la prise de décision ? » a analysé également les paramètres diagnostiques des données biologiques.

Cette partie de l'analyse a reposé principalement sur les résultats de la méta-analyse d'Anderson (51) et sur l'étude de Cardall *et al.* (54) analysées par ailleurs dans ce travail et également sur les données les résultats de 2 études rétrospectives de niveau 3. La principale conclusion qui en ressort est celle décrite précédemment en § 5.2.1.

« Chez les patients avec suspicion d'appendicite aiguë, « utilisez » l'examen clinique (signes et symptômes) pour stratifier le risque et orienter la décision concernant les examens complémentaires (pas d'examens, examens biologiques, et/ou imagerie) et celle concernant la prise en charge ultérieure du patient (sortie, observation et/ou consultation chirurgicale)» niveau B

► Conclusion

Au travers de l'analyse des articles relatifs à la valeur des signes inflammatoires dans le diagnostic de l'appendicite, il apparaît qu'aucun paramètre biologique de l'inflammation n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.

5.2.5 Valeur diagnostique des signes combinés

En 1996, l'ANDEM (38) en collaboration avec les professionnels, a pu dégager un consensus sur les pratiques recommandables en matière de diagnostic et de prise en charge de l'appendicite. Les mêmes recommandations ont été publiées par la SNFGE en 2001 (10).

Concernant le diagnostic clinico-biologique, il a été identifié trois facteurs qui lorsqu'ils sont tous les trois absents permettaient d'exclure le diagnostic d'appendicite grave et de surseoir à l'appendicectomie :

- l'absence de fièvre supérieure à 38°;
- l'absence de défense localisée à la fosse iliaque droite ;
- l'absence de leucocytose supérieure à $10 \cdot 10^9$ cellules /L.

► Analyse des articles

Depuis 1996, très peu d'articles publiés dans la littérature portent sur l'évaluation de la performance diagnostique des signes combinés. Seuls trois articles, ont étudié en partie la combinaison de certaines variables. Il s'agit de la méta-analyse d'Anderson (51) de l'étude prospective d'Ortega-Deballon et al. (46), et de l'étude de Laméris et al. (47). Les caractéristiques des patients de même que la qualité méthodologique de ces études sont décrites dans le § 5.2.2. Hormis la méta-analyse, aucune étude n'évalue plus de 2 variables à la fois.

Méta-analyse Anderson (51)

Dans la méta-analyse d'Anderson (51), l'évaluation de la valeur diagnostique des variables combinées ne fait pas l'objet de la méta-analyse. Chaque étude analysée a évalué une combinaison différente ne permettant pas la compilation des résultats. Les valeurs sont donc décrites dans cette étude à titre narratif. Les principaux résultats décrits sont ceux de cinq études originales qui ont analysé la combinaison des variables inflammatoires (tableau 7).

A la lecture des résultats les combinaisons qui sont retenues eu égard à leur capacité discriminante (cf. tableau 6), sont une hyperleucocytose supérieure à $10 \cdot 10^9$ cellules/L associée à une CRP supérieure à 8 mg/L ou à une défense abdominale.

Comparativement aux données des variables individuelles (cf. tableaux 1, 2, 5), l'association d'une hyperleucocytose à $10 \cdot 10^9$ cellules/L et une concentration de CRP supérieure à 8 mg/L montre une capacité discriminante et des rapports de vraisemblance largement supérieures avec une AUC ROC qui atteint une valeur de 0,96 IC 95%, [0,92-1,0] un RVP de 23,32 [6,87-84,79] et un RVN de 0,03 IC 95% [0,00-0,14].

Avec l'association d'une hyperleucocytose supérieure $10 \cdot 10^9$ cellules /L et d'une défense abdominale l'AUC ROC atteint une valeur de 0,84 IC 95%, [0,80-0,88] les rapports de vraisemblance des valeurs de 11,34 IC 95%, [6,65-19,56] pour le RVP et 0,14 IC 95%, [0,08-0,24] pour le RVN.

Les études prospectives

Dans l'étude d'Ortega-Deballon et al. (46), la performance diagnostique de la CRP (>6mg /L) est augmentée quand elle est combinée à un nombre de leucocytes supérieur à $10 \cdot 10^9$ cellules/L, avec une AUC ROC de 0,87 ($p < 0,0005$) et des valeurs prédictives positive de 93,2% et négative de 92,3% ($p < 0,0005$). Les RVP et RVP calculés à partir des valeurs moyennes de sensibilité et spécificité sont respectivement de 1,36 et 0,38.

Dans l'étude de Laméris et al. (47), la combinaison d'une leucocytose supérieure à $10 \cdot 10^9$ cellules/L à une concentration de CRP supérieure à 12 mg/L présente un RVP de 1,4 IC 95% [1,2-1,7] et un RVN de 0,09 IC 95%. [0,03-0,3].

Comparés aux valeurs rapportées dans la méta-analyse d'Anderson, les RVP des études (cf. tableau 7) sont nettement inférieurs. Les auteurs ont expliqué cette différence par la faible qualité

des études de la méta-analyse qui souffre en particulier d'une grande hétérogénéité des patients inclus.

Tableau 7. Variables combinées, capacité discriminante (AUC) et performance diagnostique (RVP et RVN)

Auteur, année, référence,	Association des signes	AUC ROC [IC95%]	RVP [IC95%]	RVN [IC95%]
Anderson, 2004 (51)	GB>10x10 ⁹ + CRP>8mg/L	0,96 [0,92-1,00]	23,32 [6,87-84,79]	0,03 [0,00-0,14]
	GB>10x10 ⁹ + CRP>12 mg/L+ PNN >70%	0,79 [0,74-0,84]	20,85 [5,47-80,27]	0,03 [0,01-0,16]
	GB>10x10 ⁹ + CRP>8mg/L+IL6>60ng/L	0,87 [0,80-0,94]	16,96 [3,08-98,66]	0,03 [0,01-0,16]
	Défense + GB>10x10 ⁹	0,84 [0,80-0,88]	11,34 [6,65-19,56]	0,14 [0,08-0,24]
	GB>10x10 ⁹ + CRP>12 mg/L	0,85 [0,80-0,90]	8,22 [4,7-14,38]	0,05 [0,01-0,18]
Laméris <i>et al.</i> , 2009 (47)	GB>10x10 ⁹ + CRP>12 mg/L	NR	1,4 [1,2-1,7]	0,09 [0,03-0,3]
Ortega-Deballon <i>et al.</i> , 2008 (46)	GB>10x10 ⁹ + CRP>6mg/L	0,87 [NR]	1,36*	0,38*

GB : (globules blancs) leucocytes, **PNN** : Polynucléaires neutrophiles, **NR** : non renseigné, **RVP** : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients.*calculés à partir des valeurs de Se et Sp

Par ailleurs, dans l'étude de Laméris *et al.* (47), il a été recherché les profils correspondant aux différentes combinaisons possibles des signes relatifs à une appendicite dite « Classique » et qui sont : une douleur migrante, une sensibilité en FID et une défense abdominale avec ou sans augmentation de leucocytes et de CRP, chez des patients âgés de moins de 30 ans et de plus 50 ans. Il en est ressorti une grande variabilité de présentations clinico-biologique, soit 23 profils différents.

Sur le tableau 8 sont présentés les résultats obtenus sur les 7 profils qui montrent un taux d'appendicite supérieur ou égal à 75%, la fréquence relative de ces 7 différents profils dans la population de l'étude est faible. La combinaison la plus fréquente est la douleur en FID avec la défense abdominale (9% des femmes et 13% des hommes). La deuxième est une douleur en FID associée à un âge inférieur à 30 ans (16% de femmes et 13% d'hommes). Les autres profils présentent tous une fréquence inférieure à 3%.

La forme clinique dite « classique » qui associe douleur migrante, douleur en FID et défense abdominale n'est retrouvée que chez 6% des patients de l'étude, 3% d'hommes et 3% de femmes. Le taux d'appendicite retrouvée dans cette catégorie est de 100% chez les hommes et 46% chez les femmes.

Cette étude montre qu'en fonction de la combinaison des éléments cliniques, la précision diagnostique est différente, elle souligne également une plus grande difficulté de diagnostic différentiel chez la femme.

Au travers de ces résultats, il est noté la rareté des combinaisons clinico-biologiques présentant une bonne performance diagnostique. De ce constat, ressort la capacité limitée des scores clinico-biologiques car même s'ils reposent sur les éléments cliniques et biologiques les plus discriminants, la faible fréquence des situations en limitent l'usage. L'imagerie pourrait dans ce contexte être utile chez une large proportion de la population (47).

Dans un autre travail publié par l'ACR en 2006 (55), il a été recherché les critères cliniques de prédiction d'une appendicite afin optimiser l'utilisation du scanner. Pour cela l'ACR s'est appuyé principalement sur l'étude de Roth et al. (56), une étude transversale de 164 patients dont l'objectif était de déterminer la valeur prédictive des paramètres cliniques chez les patients souffrant de douleurs abdominales non spécifiques.

Les résultats des examens de scanner ont été corrélées avec les données cliniques et le diagnostic final afin d'évaluer leur valeur prédictive positive.

L'étude rapporte qu'un diagnostic d'appendicite est corrélé avec : une numération leucocytaire élevée ($> 11,5 \cdot 10^9$ cellules/L $p = 0,002$), le sexe masculin ($p = 0,001$) et le jeune âge (moins de 25 ans $p = 0,002$). La combinaison de l'âge, du sexe et du nombre de leucocytes permet une règle de prédiction avec aire sous la courbe opérateur récepteur (ROC) de 0,92.

Cette étude descriptive a été classée par l'ACR en niveau d'évidence 2.

► Conclusion

Au regard des données de la littérature, il apparaît que devant une douleur abdominale et une suspicion d'appendicite :

- l'association des signes cliniques tels : une douleur migrante, une douleur en FID et une défense abdominale chez un sujet de sexe masculin est très évocatrice d'une appendicite aiguë ;
- l'absence concomitante d'une défense en fosse iliaque droite, d'une hyperleucocytose supérieure à $10 \cdot 10^9$ cellules/L et d'une CRP supérieure à 8 mg/L rend le diagnostic d'appendicite peu probable ;
- l'association d'une défense en fosse iliaque droite, d'une hyperleucocytose supérieure à $10 \cdot 10^9$ cellules/L et d'une CRP supérieure à 8 mg/L est très évocatrice d'une appendicite chez l'homme jeune.

Tableau 8. Les différents profils clinico-biologiques permettant d'obtenir une prévalence d'appendicites $\geq 75\%$ (47)

Profils	Symptômes et signes cliniques			Biologie		% par rapport à l'ensemble des patients		Taux d'appendicites	
	Douleur en FID	Douleur migrante	Défense abdominale	Leucocytes $>10 \times 10^9/L$	CRP >12 mg/L	femmes	hommes	femmes	hommes
Profil 1	+	+	+			3%	3%	46	100%
Profil 2	+		+			9%	13%	50%	89%
Profil 3 âge >50 ans	+	+				2%	3%	57%	92%
Profil 4 âge >50 ans	+	+		+	+	1%	2%	75%	89%
Profil 5 âge <30 ans	+	+				5%	5%	39%	79%
Profil 6 âge <30 ans	+					16%	13%	45%	75%
Profil 7 âge <30 ans	+	+		+	+	2%	2%	60%	75%

5.2.6 Les scores clinico-biologiques

Au regard des éléments décrits précédemment, la combinaison d'informations cliniques et de marqueurs biologiques de l'inflammation montre un pouvoir discriminant supérieur à celui des variables considérées séparément.

Sur la base de ce principe, plusieurs algorithmes ont été proposés pour évaluer en fonction d'un système de scores, la probabilité d'appendicite aiguë.

► Score d'Alvarado

Le score d'Alvarado également connu sous l'acronyme de MANTRELS, est le plus connu et le plus décrit dans la littérature. Il est construit sur une échelle de 10 points et repose sur les éléments cliniques et biologiques suivants :

Signes et symptômes	Score
Douleur migrante vers la fosse iliaque droite (FID)	1
Anorexie	1
Nausée/vomissements	1
Température $\geq 37,3^{\circ}\text{C}$	1
Défense en FID	2
Défense à la percussion	1
hyperleucocytose	2
Taux neutrophiles $>75\%$	1
Total	10

D'après Brigand et al (25)

Trois niveaux de scores, auxquels sont attribuées les probabilités d'appendicite suivantes : 30 % pour [1-4], 66% pour [5-6], 93% pour [7-10] ont été identifiés (57).

Score de 1-4	faible	retour du patient à domicile
Score 5-6	modéré	observation
Score 7-10	élevé	chirurgie

Le principal biais méthodologique de ce score est qu'il a été élaboré de manière rétrospective à partir d'une population de patients opérés, il n'a pas été construit sur les données comparables à celles de la population à laquelle il devrait s'adresser (patients avec suspicion d'appendicite). La performance diagnostique et la pertinence d'utilisation de ce score en sont par conséquent discutées

Revue systématique (57)

Dans une revue récente systématique avec méta-analyse, Ohle et al. (57) ont analysé 42 études (2842 patients) qui évaluaient la validité du score d'Alvarado, l'objectif étant :

- de définir la précision et la capacité discriminante les deux valeurs du score qui sont le plus souvent utilisées pour prendre une décision : 5 pour l'observation et 7 pour la chirurgie ;
- de calibrer la précision diagnostique en fonction des 3 différents niveaux de scores : faible (1 à 4), moyen (5 à 6) et élevé (7 à 10) en stratifiant sur l'âge et le sexe : homme, femmes et enfants (âge <18 ans).

La calibration de la précision diagnostique a été évaluée en termes de risque relatif d'appendicite défini comme : le nombre d'appendicites prédites par le score sur le nombre réel d'appendicites (RR IC 95%) étant considéré qu'un $RR > 1$ signifie une surestimation du nombre d'appendicites par le score, qu'un $RR < 1$ signifie une sous-estimation et qu'un $RR = 1$ est témoin d'une parfaite correspondance entre le nombre prédit et le nombre réel. Les caractéristiques de cette étude sont décrits tableau 8 (annexe 2).

Conformément à ce qui est décrit précédemment, la principale limite méthodologique de cette revue systématique est l'hétérogénéité des études analysées. Une grande variabilité est retrouvée notamment en termes de prévalence d'appendicite, de sex-ratio, de résultats des examens cliniques. L'auteur a évoqué une absence de données de suivi des patients non opérés dans la majorité des études analysées ce qui constitue une source de biais susceptible d'augmenter artificiellement les valeurs de sensibilité et de spécificité.

Les résultats de cette étude (cf. tableau 9) montrent, quelle que soit la population étudiée, qu'un score de 5 a une sensibilité élevée pour orienter le patient vers l'observation et exclure une appendicite dans l'immédiat. En revanche, un score de 7, présente une faible valeur diagnostique (sensibilité et spécificité modérées) qui ne permet pas de confirmer ou d'exclure avec certitude le diagnostic d'appendicite.

Tableau 9. Valeurs agrégées de sensibilité et de spécificité aux valeurs de score 5 et 7(57)

	Score 5 (observation)		Score 7 (chirurgie)	
	Sensibilité [IC 95%]	Spécificité [IC 95%]	Sensibilité [IC 95%]	Spécificité [IC 95%]
Population générale	0,99 [0,97-0,99]	0,43 [0,36-0,51]	0,82 [0,76-0,86]	0,81 [0,76-0,85]
Hommes	0,96 [0,88-0,99]	0,34 [0,24-0,47]	0,88 [0,75-0,95]	0,57 [0,40-0,73]
femmes	0,99 [0,92-0,99]	0,35 [0,14-0,64]	0,86 [0,78-0,92]	0,73 [0,58-0,84]
enfants	0,99 [0,83-1,00]	0,57 [0,41-0,72]	0,87 [0,76-0,93]	0,76 [0,55-0,89]

Le tableau 9 montre les résultats des rapports de risque évalué en fonction des niveaux de scores et en fonction de l'âge et du sexe. Quelle que soit la catégorie de risque, le score semble relativement mieux équilibré chez l'homme avec un RR maximal de 1,09. Chez la femme, une surestimation du risque est observée à tous les niveaux de scores avec un RR pouvant atteindre jusqu'à 5,35 (tableau 10).

Tableau 10. Rapports de risque en fonction du niveau de score et en fonction de l'âge et du sexe (57)

	Rapport de risque [IC 95%]		
	Score 1 à 4	Score 5 à 6	Score 7 à 10
Hommes	1,06 [0,87-1,28]	1,09 [0,86-1,37]	1,02 [0,97-1,08]
femmes	5,35 [2,17-13,9]	1,82 [1,20-2,78]	1,14 [1,04-1,25]
enfants	5,03 [0,52-48,82]	1,81 [1,13-2,89]	1,13 [1,01-1,27]

Les auteurs concluent que le score d'Alvarado est mieux calibré chez l'homme que chez la femme. C'est un bon outil diagnostique pour détecter les patients qui doivent être suivis en observation (score >5). En revanche, il ne permet pas de trancher pour la chirurgie chez les patients qui ont un score de 7, des examens complémentaires sont nécessaires pour confirmer le diagnostic d'appendicite.

Recommandations de l'ACEP

L'évaluation de l'utilité du score Alvarado a été analysé dans le travail de l'ACEP (53) sur 4 études dont 3 rétrospectives et une prospective, toutes de faible niveau de preuves (niveau 3). Les résultats des études sont très hétérogènes et controversés. Le travail de l'ACEP ne mentionne pas de recommandation particulière quant à l'utilisation du score d'Alvarado.

Mais, le texte de la recommandation (voir § 5.1.2) rejoint dans la teneur les conclusions de la revue systématique d'Ohle et al. (57) en attribuant aux données clinico-biologiques la seule valeur d'aide à la stratification du risque.

► Score d'Andersson

Une seule étude prospective de 545 patients a décrit ce score dans la littérature (24). A partir d'un échantillon de 316 patients admis pour suspicion d'appendicite, ont été recueillies plusieurs variables diagnostiques identifiées dans la littérature. Ces variables ont été analysées selon un modèle de régression linéaire pondéré qui a permis d'identifier 8 variables indépendantes qui ont été ensuite utilisées pour la construction du score sur une échelle de 0 à 12 points :

Signes et symptômes		Score
Vomissements		1
Douleur en FID		1
Douleur à la décompression ou défense abdominale	Léger	1
	Moyen	2
	Fort	3
Fièvre >38,5°C		1
Taux de PNN	70-84%	1
	≥ 85%	2
Leucocytes	10-14,9 10 ⁹ /L	1
	≥ 15 10 ⁹ /L	2
CRP	10-49 g/L	1
	≥ 50 g/L	2
Somme (0 à 12)		

Trois zones de probabilité diagnostique sont définies :

Score <5	Faible risque	retour du patient à domicile
Score 5-8	doute	observation
Score >8	élevé	chirurgie

Dans cette étude, ce score est validé sur un échantillon 229 patients chez qui a été réalisé également le score d'Alvarado à une fin de comparaison.

Les caractéristiques des populations entre les deux échantillons de construction et de validation du score sont comparables en termes de prévalence d'appendicite respectivement 36 vs 33% et de sex-ratio, de taux d'appendicites négatives 10% vs 11%. L'âge moyen des patients est de 25,9 ans dans l'échantillon de construction vs 23,4 (p=0,022) dans l'échantillon de validation.

Les résultats de l'étude montrent que le score d'Andersson peut classer correctement 73% des patients dans le groupe à faible risque (score <5) et 67% des patients dans le groupe à risque élevé (score >8).

La comparaison des deux scores montre en termes de capacité discriminante une AUC ROC pour le score d'Andersson de 0,93 vs 0,88 ($p=0,0007$) pour le score d'Alvarado. Concernant les appendicites perforées, l'AUC ROC est respectivement de 0,97 et 0,92 ($p=0,0027$).

Sur les deux points les plus culminants (point de sensibilité élevée et point de spécificité élevée), cette étude montre que le score d'Andersson permet une classification de 63 % des patients à faible risque et à risque élevé avec une précision de 0,97 vs 52% pour le score Alvarado précision de 0,98 ($p=0,023$). En définitive, 37% des patients sont mal classés avec le score Andersson vs 48% avec le score d'Alvarado.

Selon les auteurs, la meilleure performance diagnostique constatée avec le score d'Andersson pourrait s'expliquer par l'amélioration des points suivants : le recueil prospectif de variables diagnostiques indépendantes et le remplacement de variables subjectives telles la nausée, l'anorexie et la migration de la douleur par des variables plus objectives comme la concentration de CRP.

Le faible effectif de cette étude et le nombre de patients exclus pour manque de données (27%) fragilisent la validité des résultats. La validation de ce score nécessite la réalisation d'une étude prospective multicentrique avec un échantillon de patients plus conséquent.

► Conclusion

Au regard de la faible qualité des données publiées, il est difficile de valider la fiabilité des ces deux scores.

Le score d'Alvarado semble présenter un intérêt relatif «chez l'homme » pour identifier les situations dans lesquelles les patients peuvent être renvoyés à domicile ou suivis en observation. Ce n'est pas un outil fiable pour identifier les patients qui doivent être opérés. Il n'est pas non plus fiable chez la femme chez qui il existe une diversité de diagnostics différentiels.

Le score d'Andersson semble présenter un intérêt supplémentaire par rapport au score d'Alvarado du fait de l'intégration de la concentration en CRP en tant que critères d'évaluation, mais le faible niveau de preuve des données publiées impose une confirmation de cette hypothèse dans le cadre d'une étude prospective méthodologiquement fiable.

5.3 Performance diagnostique de l'imagerie

L'imagerie est utilisée pour augmenter la précision diagnostique d'appendicite en particulier quand le tableau clinique n'est pas concluant.

Dans ce travail, il s'agit d'évaluer la performance diagnostique des deux méthodes les plus communément utilisées : l'échographie et le scanner.

Après sélection de la littérature identifiée, les documents suivants ont été retenus :

- 3 méta-analyses (48,58,59) qui ont évalué simultanément le scanner et l'échographie ;
- Une méta-analyse qui a évalué la performance diagnostique du scanner réalisé sans produit de contraste (60) ;
- 2 recommandations (36,53).

Onze études prospectives dont 6 études contrôlées randomisées et 4 méta-analyses ont été exclues de l'analyse pour les raisons suivantes :

- Pour les études :
 - 7 articles déjà analysées dans les méta-analyses et/ou les recommandations et dont les conclusions sont compatibles avec ces documents de synthèse ;(61-67) ;
 - 1 études dont la validité méthodologique est jugée discutable au regard des critères de la grille QUADAS-2, (68) ;
 - 1 étude dont les résultats ne permettent pas l'estimation de la Se, Sp et des RVP et RVN(69) ;

- ▶ 1 étude qui analyse la performance diagnostique du scanner avec une modalité de contraste oral (39), Cette modalité n'est pas conforme à la pratique française qui ne préconise pas l'opacification digestive par voie haute (31).
- Pour les méta-analyses et revues systématiques :
 - ▶ biais de sélection géographique, la méta-analyse de Yu et al. a inclus exclusivement les études publiées en coréen (70) ;
 - ▶ 4 méta-analyses de faible qualité méthodologique selon les critères de la grille d'analyse AMSTAR13, (71-74) ;
 - ▶ 1 méta-analyse dont l'objectif n'est pas l'évaluation de la performance diagnostique (75).

5.3.1 Valeur diagnostique de l'échographie

Trois méta-analyses (48,58,59) ont évalué la performance diagnostique de l'échographie. En raison de la variabilité des objectifs des études, des caractéristiques des patients inclus, des techniques d'échographie utilisées, les résultats de chaque étude ont été analysés de manière individuelle.

▶ Méta-analyse de van Randen et al. (58)

L'objectif principal de la méta-analyse de van Randen (58) était de comparer directement chez le même patient la valeur diagnostique de l'échographie à compression progressive vs celle du scanner chez une population se présentant avec une douleur abdominale aiguë.

Le deuxième objectif était d'estimer la probabilité post-test du scanner et de l'échographie chez 3 sous groupes de patients avec des prévalences différentes : patients avec suspicion d'appendicite, patients avec douleur abdominale aiguë (DAA) et patients avec une douleur en FID.

Le test de référence est l'examen anatomopathologique pour les patients opérés et le suivi clinique chez les patients non opérés. Six études prospectives sont analysées ce qui représente un total de 671 patients.

Caractéristiques des patients

L'âge moyen des patients dans les études est compris entre 26 et 38 ans avec un sex-ratio (femmes/hommes) très variable allant de 0,32 à 2,24. La majorité des patients sont admis pour suspicion d'appendicite (540 sur 671). En fonction des études, la prévalence d'appendicite varie de 13 à 77%. Le suivi moyen des patients non opérés varie de 1 à 13 mois ;

Techniques de l'échographie

Les caractéristiques techniques des sondes à ultrasons utilisées sont très variables : linéaires, courbes, curvilinéaires ou convexe avec une fréquence de sonde variant de 3,5 à 12 MHz. Dans une même étude, plusieurs technologies pouvaient être utilisées (3 études sur 6).

La région de l'exploration échographique n'est pas renseignée de même que les critères échographiques de l'appendicite. L'expérience des radiologues décrite dans les études originales est très variable : de radiologiste expérimenté avec plus de 10 ans d'expérience à radiologiste résidents avec 3 à 7 ans d'expérience à radiologiste assistant.

Analyse de l'hétérogénéité

L'analyse de l'hétérogénéité a été recherchée et l'estimation de la variabilité inter-études des valeurs moyennes de sensibilité et spécificité est réalisée au moyen d'un modèle bivarié de régression linéaire à effets aléatoires.

¹³ Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, *et al.* Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10

Résultats de performance diagnostique

Cette étude rapporte les résultats sensibilité de 0,78, une spécificité de 0,83 et des rapports de vraisemblance positifs et négatifs respectivement de 4,5 et 0,27 (cf. tableau 11).

Variabilité de la probabilité post-test en fonction de la prévalence

Sur la base de 17 études (5093 patients), la prévalence de l'appendicite en fonction des populations a été établie par les auteurs comme suit :

- patients avec douleur en FID : probabilité pré-test 37% ;
- patients avec douleur abdominale aiguë (DAA) : probabilité pré-test 6% ;
- patients avec DAA et âgés de plus de 65 ans probabilité pré-test 3%.

Sur l'ensemble de la population de cette méta-analyse dont la prévalence moyenne d'appendicite est de 50% [13%-77%], la probabilité post test pour un test positif est de 82% et pour un test négatif de 21%.

En fonction de la situation clinique, cette probabilité post-test variait :

- pour un test positif, elle variait de 73 % pour les patients avec douleur en FID (probabilité pré-test 37%), à 22% pour les patients avec DAA à 12% pour les patients avec DAA et âgés de plus de 65 ans ;
- pour un test négatif, elle variait de 14 % pour les patients avec douleur en FID à 2% pour patients avec DAA à 1% pour les patients avec DAA et âgés de plus de 65 ans.

Les principales limites méthodologiques

Les principales limites méthodologiques relevées dans cette étude sont l'hétérogénéité des patients des études, du fait de la sélection sous le critère générique de « douleur abdominale aiguë » et d'un sex-ratio très différent entre les études. Cette hétérogénéité transparaît dans le taux de prévalence entre les différentes études. De plus, les données de suivi clinique pour les patients non-opérés ne sont pas rapportées, d'où le risque de surestimation de l'intérêt du test.

Conformément à ce qui est décrit en chapitre 5.1.2, cette étude souffre intrinsèquement d'un biais de vérification, d'un biais d'absence d'insu car le test de référence est réalisé en connaissance des résultats de l'imagerie. Les biais de publication n'ont pas été vérifiés.

Dans cette étude, le poids et l'IMC des patients n'ont pas été rapportés. Ces données sont particulièrement importantes pour les patients obèses dont l'examen échographique est souvent non contributif.

Au regard des données présentées ci-dessus, l'échographie semble présenter une performance diagnostique modérée, avec un rapport de vraisemblance négatif qui montre qu'une échographie négative ne permet pas d'éliminer catégoriquement le diagnostic d'appendicite. Cette étude montre par ailleurs que la performance diagnostique de l'échographie est intimement liée à la prévalence de l'appendicite dans la population étudiée.

► Méta-analyse de Doria (76)

L'objectif principal de la méta-analyse de Doria (76) était d'évaluer la performance diagnostique de l'échographie et du scanner pour le diagnostic de l'appendicite chez les populations adulte et pédiatrique. Le test de référence adopté était l'examen anatomopathologique pour les patients opérés et un suivi clinique d'au moins une semaine chez les patients non opérés.

Douze études prospectives et une étude rétrospective ont été analysées, représentant un total 1579 patients.

Caractéristiques des patients

L'âge moyen des patients dans les études était compris entre 20 et 46 ans avec un sex-ratio (femmes/hommes) très variable allant de 0,52 à 5,25 (calculé à partir des données). La présentation clinique et le degré de sévérité chez les patients inclus dans les études n'ont pas été décrits. La prévalence moyenne sur l'ensemble de la population de la méta-analyse était de 48%,

Techniques de l'échographie

La majorité des études ont utilisé des sondes à ultrasons linéaires, avec une fréquence d'au moins 5 MHz. Les critères échographiques de visualisation d'un appendice inflammatoire étaient : un diamètre supérieur à 6 mm et un appendice non compressible à l'échographie. En cas de non visualisation, la présence de signes inflammatoires. La région de l'exploration échographique n'était pas renseignée de même que le niveau d'expérience des radiologues n'est pas décrit

Analyse de l'hétérogénéité

Dans cette étude, après analyse de la sensibilité, les études qui apportaient l'hétérogénéité ont été exclues. Une analyse de méta-régression¹⁴ a été réalisée pour étudier la relation entre les caractéristiques des études (le schéma de l'étude (prospective vs rétrospective), l'année de publication et le continent d'origine) et les valeurs de sensibilité et de spécificité. Aucune hétérogénéité n'a pu être mise en évidence (pour la sensibilité $\chi^2=1,75$ ($p=0,19$), spécificité $\chi^2=1,26$ ($p=0,26$). Les sensibilités et spécificités ont par conséquent été calculées selon un modèle à effets fixes

Résultats de performance diagnostique

Les valeurs de sensibilité et de spécificité estimées par la méta-analyse sont respectivement de 83% IC 95% [78-87] et de 94% IC 95% [92-95]. Les valeurs de rapports de vraisemblances calculées à partir des valeurs de sensibilité et de spécificité estimées par la méta-analyse sont de 13,8 pour le RVP et de 0,18 pour le RVN (cf. Tableau 10).

Les principales limites méthodologiques

Les principales limites méthodologiques de cette étude sont inhérentes à la qualité des études originales, celles-ci souffrent principalement des biais suivants :

- d'un biais de vérification ;
- biais d'absence d'insu, le test de référence n'a pas été réalisé en insu des résultats d'imagerie ;
- d'une absence de description de la gravité des signes d'appendicite ce qui ne permet pas d'appréhender la probabilité pré-test et donc d'apprécier le biais de spectre.

Les auteurs rapportent également une absence de description des données sur le poids et l'IMC des patients. Facteurs particulièrement importants pour les patients obèses dont l'examen échographique est souvent non conclusif.

Pour les résultats équivoques, les auteurs rapportent une difficulté à définir un seuil de positivité. L'absence de visualisation de l'appendice est considérée dans certaines études, comme un résultat négatif, dans d'autres ce résultat équivoque pourrait également évoquer une appendicite perforée.

Au regard des données présentées ci-dessus l'échographie semble présenter une performance diagnostique modérée, notamment pour la sensibilité, ce qui ne permet pas d'éliminer une appendicite en cas d'échographie négative.

Toutefois, le faible niveau de preuves des études originales confère à cette étude une évidence relative.

¹⁴ Il s'agit d'une technique de méta-analyse multivariée, comme la régression logistique, utilisée pour étudier, dans une synthèse méthodique, la relation entre les caractéristiques des études (par exemple le secret d'attribution, le risque de sortie, le calendrier de l'intervention) et les résultats des études (l'ampleur de l'effet dans chaque étude).

► Méta-analyse de Terasawa et al. (48)

L'objectif principal de la méta-analyse de Terasawa et al. (48) était de définir la précision diagnostique de la tomodensitométrie et de l'échographie pour le diagnostic de l'appendicite chez les adultes et les adolescents avec suspicion d'appendicite aiguë. Le test de référence était l'examen anatomopathologique pour les patients opérés et le suivi clinique chez les patients non opérés.

Quatorze études prospectives répondant aux critères de sélection ont été analysées, elles totalisaient en tout 1516 patients.

Caractéristiques des patients

L'âge moyen des patients dans les études était compris entre 20 et 46 ans avec un sex-ratio (femmes/hommes) moyen de 1,22. La majorité des patients (85%) présentaient un tableau de « suspicion d'appendicite » et 14,5% avaient des présentations cliniques « atypiques ».

En fonction des études, la prévalence d'appendicite variait de 13 à 75% avec une prévalence moyenne autour de 50%. Le taux moyen d'appendicites négatives variait de 14 à 50 % avec une valeur moyenne de 36% (valeurs disponibles uniquement dans 10 études).

Techniques de l'échographie

Les caractéristiques techniques des sondes à ultrasons utilisées étaient très variables : linéaires, courbes, curvilinéaires ou convexe avec une fréquence de sonde variant de 2,5 à 9,0MHz. Dans une même étude, plusieurs technologies et plusieurs fréquences pouvaient être utilisées. La région de l'exploration échographique n'était pas renseignée de même que le niveau d'expérience des radiologues n'était pas décrit.

Le seuil de positivité à l'échographie correspondait à un diamètre de l'appendice supérieur à 6 mm, d'autres critères ont également été utilisés dans les études tels : un appendice non compressif et apéristaltique, un œdème péri-appendiculaire, un stercolithe ou un abcès appendiculaire.

Analyse de l'hétérogénéité

Dans cette étude il n'a pas été tenu compte de l'hétérogénéité des études pour estimer les valeurs de sensibilité et spécificité dans la méta-analyse. En revanche les auteurs ont cherché à évaluer le degré d'hétérogénéité des rapports de vraisemblance des études en utilisant le test Q de Cochrane et I² (indice de Higgins et al.¹⁵ qui varie de 0 à 100 : une valeur I² < 25 indique une hétérogénéité faible, des valeurs comprises entre 25 et 50 une hétérogénéité modérée et une valeur > 50 une hétérogénéité importante).

Une hétérogénéité significativement élevée a été retrouvée avec les deux tests, pour les RVP I²=85% (IC 95% : 76-90%) et Q= 86,33 (p<0,001) et pour les RVN I²=75% (IC 95% : 59-85%) et Q= 52,75 (p<0,001). Un modèle d'effets aléatoire a été utilisé pour l'analyse des résultats.

Résultats de performance diagnostique

Cette étude rapporte les résultats sensibilité de 0,86 IC 95% [0,83-0,88], une spécificité de 0,81 IC 95% [0,78-0,84] et des rapports de vraisemblance positifs et négatifs respectivement de 5,8 IC 95% [3,5-9,5] et 0,19 [0,13-0,27] (cf. Tableau 10).

Les principales limites méthodologiques

Une hétérogénéité significativement élevée a été observée entre les études, probablement du fait de la variabilité de présentation clinique et du degré de sévérité chez les patients, de la variabilité des pratiques, des sites d'examen différents et de l'expérience des opérateurs.

¹⁵ Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60

Les principales limites méthodologiques de cette étude sont inhérentes à la faible qualité méthodologique des études originales qui souffrent principalement d'un biais de vérification différentiel, d'un biais d'absence d'insu pour le test de référence (anatomopathologie).

Au regard des données de la méta-analyse, l'échographie semble présenter une performance diagnostique modérée mais toutes les limites citées plus haut sont susceptibles de biaiser ces valeurs, ce qui confère à cette étude une évidence relative.

► Conclusion

A l'analyse de la littérature, il apparaît qu'en règle générale, les études originales présentent des biais majeurs susceptibles d'induire une surestimation des valeurs de la performance diagnostique de l'échographie. Une grande hétérogénéité est également constatée, la performance diagnostique de l'échographie est variable d'une étude à l'autre, elle est patient-dépendant et opérateur-dépendant.

L'intérêt diagnostique de l'échographie paraît modérée, elle semble plus utile pour confirmer le diagnostic d'appendicite que pour l'exclure.

Tableau 11. Valeur diagnostique de l'échographie pour le diagnostic de l'appendicite chez l'adulte, résultats de méta-analyses

Auteur, Année, référence	Nombre d'études, type	Nombre de patients	Age moyen	Présentation clinique	Prévalence (bornes extrêmes)	Sensibilité [IC 95%]	Spécificité [IC 95%]	RVP [IC 95%]	RVN [IC 95%]
Terasawa et al., 2004 (48)	14 prospectives	1516	de 31 à 38 ans	1296 suspicions d'appendicite 220 atypiques	50% [13%-75%]	0,86 [0,83-0,88]	0,81 [0,78-0,84]	5,8 [3,5-9,5]	0,19 [0,13-0,27]
Doria, 2006 (76)	16 prospectives et 1 rétrospective	1816	de 20 à 49 ans	NR	48% [NR]	0,83 [0,78-0,87]	0,94 [0,92-0,95]	13,8*	0,18*
van Randen et al., 2008 (58)	6 prospectives	671	de 26 à 38 ans	540 suspicions d'appendicite 131 atypiques	50% [13%-77%]	0,78 [0,67-0,86]	0,83 [0,76-0,88]	4,5 [3,03-6,68]	0,27 [0,17-0,43]

*Calculé à partir des valeurs de Se et Sp ;NR : non renseigné **RVP** : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients.

Tableau 12. Valeur diagnostique du scanner pour le diagnostic de l'appendicite chez l'adulte, résultats de méta-analyses

Auteur, Année, référence	Nombre d'études, type	Nombre de patients	Age moyen	Présentation clinique	Prévalence [bornes extrêmes)	Sensibilité [IC 95%]	Spécificité [IC 95%]	RVP [IC 95%]	RVN [IC 95%]
Terasawa et al., 2004 (48)	12 prospectives	1172	de 30 à 38 ans	668 atypiques 504 suspicions d'appendicite	45% [13-78]	0,94 [0,91-0,95]	0,95 [0,93-0,96]	13,3 [9,9-17,9]	0,09 [0,07-0,12]
Doria, 2006 (76)	15 prospectives et 4 rétrospectives	2394	de 20 à 49 ans	NR	40% [NR]	0,94 [0,92-0,95]	0,94 [0,94-0,96]	15,66*	0,06*
van Randen et al., 2008 (58)	6 prospectives	671	de 26 à 38 ans	540 suspicions d'appendicite 131 atypiques	13%-77%	0,91 [0,84-0,95]	0,90 [0,85-0,94]	9,3 [6,86-12,6]	0,1 [0,06-0,17]
Hlibczuk et al., 2010 (60)	7 prospectives	1060	NR*	1060 Douleur abdominale aiguë avec suspicion d'appendicite	39,3 [20,1-84,5]	0,92 [0,89-0,95]	0,96 [0,94-0,97]	24	0,08

*Calculé à partir des valeurs de Se et Sp ;NR : non renseigné **RVP** : rapport de vraisemblance positif, **RVN** : rapport de vraisemblance négatif, **IC** : intervalle de confiance, **N** : nombre de patients.**NR**

5.3.2 Valeur diagnostique du scanner

La performance diagnostique du scanner a été évaluée dans 4 méta-analyses (48,58-60), pour les 3 études dont les caractéristiques générales ont été précédemment décrites dans la partie échographie, seules les techniques de scanner et les caractéristiques des patients inclus dans les études évaluant le scanner ont été rapportées.

Conformément à ce qui a été adopté dans le paragraphe précédent, la variabilité des études en termes : d'objectifs, de caractéristiques des patients inclus et de techniques de scanner utilisées a imposé de décrire de façon individuelle les différentes études.

► Analyse des articles

Méta-analyse de Terasawa et al. (48)

Dans cette étude la performance diagnostique du scanner a été évaluée sur 12 études prospectives totalisant 1172 patients.

Caractéristiques des patients dans les études scanner

L'âge moyen des patients dans les études était compris entre 31 et 38 ans avec un sex-ratio (femmes/hommes) moyen de 1,56. La majorité des patients (57%) avaient des présentations cliniques « atypiques », 43% présentaient un tableau clinique de « suspicion d'appendicite ».

En fonction des études, la prévalence d'appendicite variait de 24 à 78% avec une prévalence moyenne autour de 45%

Techniques du scanner

Une grande hétérogénéité a été retrouvée, entre les études en termes de technique de scanner (hélicoïdal et non hélicoïdal), de site d'examen (abdomino-pelvien ou focalisé) et en termes de modalités du produit de contraste (oral, intraveineux rectal, oral plus intraveineux, ou plusieurs modalités dans une même étude).

Le critère de positivité au scanner adopté dans les études correspondait à un diamètre de l'appendice supérieur à 6 mm, d'autres critères tels que l'inflammation péri-appendiculaire, l'appendicolithe ou l'épaisseur de la paroi caecale ont également été utilisés.

Analyse de l'hétérogénéité

Aucune hétérogénéité statistiquement significative n'a été mise en évidence concernant les rapports de vraisemblance : pour les RVP $I^2=0\%$ (IC 95% : 0-58%) et $Q= 9,16$ ($p<0,001$) et pour les RVN $I^2=0\%$ (IC 95% : 0-58%) et $Q= 7,38$ ($p<0,001$). Un modèle à effets fixe a été utilisé pour la méta-analyse.

Résultats de performance diagnostique

Cette étude rapporte une sensibilité de 0,94 IC 95% [0,90-0,95], une spécificité de 0,95 IC 95% [0,93-0,96] et des rapports de vraisemblance positifs et négatifs respectivement de 13,3 IC 95% [9,9-17,9] et 0,09 IC 95% [0,07-0,13] (cf. Tableau 11).

Au regard des données présentées ci-dessus, le scanner montre une bonne performance diagnostique. Celle-ci reste relativement homogène malgré la variabilité de la présentation clinique, du degré de sévérité du tableau clinique chez les patients, des techniques, des pratiques de scanner employées et des sites d'examens différents.

Cependant, la faible qualité méthodologique des études originales et notamment la présence de biais multiples (cf. 5.3.1) susceptibles d'induire une surestimation de l'effet, confère un niveau de preuve moindre à cette étude.

Méta-analyse de van Randen et al. (58)

Pour cette étude qui comparait la performance diagnostique du scanner et de l'échographie chez les mêmes patients (671 patients), les caractéristiques des patients n'ont pas été rappelées, elles sont décrites en §5.3.1.

Techniques du scanner

Une grande hétérogénéité existait concernant : les techniques utilisées dans les différentes études (hélicoïdal monobarrette ou multibarrette, non hélicoïdal), les sites d'examen (abdominal, pelvien, abdomino-pelvien ou focalisé) et les modalités de contraste (sans contraste, oral, intraveineux rectal, oral plus intraveineux, ou plusieurs modalités dans une même étude).

Les critères de positivité du scanner étaient variables selon les études et dans une seule étude sur les six de la méta-analyse, ils n'ont pas été renseignés.

Analyse de l'hétérogénéité

L'analyse de l'hétérogénéité a été recherchée et la prise en compte de la variabilité inter-études des valeurs moyennes de sensibilité et spécificité a été réalisée au moyen d'un modèle bivarié de régression linéaire à effets aléatoires.

Résultats de performance diagnostique

Cette étude rapporte une sensibilité de 0,91 IC 95% [0,84-0,95], une spécificité de 0,90 IC 95% [0,85-0,94] et des rapports de vraisemblance positifs et négatifs respectivement de 9,29 IC 95% [6,86-12,6] et 0,10 IC 95% [0,06-0,17] (cf. Tableau 11). Le taux de faux positifs varie de 2 à 11%.

Variabilité de la probabilité post-test en fonction de la prévalence

Sur la base de 17 études (5093 patients), la prévalence de l'appendicite en fonction des populations a été établie par les auteurs comme suit :

- patients avec douleur en FID : probabilité pré-test 37% ;
- patients avec douleur abdominale aiguë (DAA) : probabilité pré-test 6% ;
- patients avec DAA et âgés de plus de 65 ans probabilité pré-test 3%.

Sur l'ensemble de la population des cette méta-analyse qui présentait une prévalence moyenne de 50%, la probabilité post test calculée était de 90% pour un test positif et de 9% pour un test négatif.

En fonction de la situation clinique, cette probabilité post-test variait :

- avec un test positif, elle variait de 85 % pour les patients avec douleur en FID à 37% pour les patients avec DAA à 22% pour les patients avec DAA et âgés de plus de 65 ans ;
- avec un test négatif elle variait de 6 % pour les patients avec douleur en FID à 1% pour patients avec DAA à 0% pour les patients avec DAA et âgés de plus de 65 ans.

A l'instar de l'étude précédente, les valeurs enregistrées sont en faveur d'une bonne performance diagnostique du scanner néanmoins les différents biais dans les études originales (biais de vérification différentiel et biais d'absence d'insu) peuvent induire une surestimation de ces résultats.

Méta-analyse de Doria (76)

Dans cette étude la performance diagnostique du scanner a été évaluée sur 5 études rétrospectives 13 prospectives totalisant 2439 patients.

Caractéristiques des patients dans les études scanner

Dans cette étude la présentation clinique de l'appendicite chez les patients n'était pas renseignée, l'âge moyen des patients dans les études était compris entre de 20 et 49 ans, le sex-ratio (F/H) variait de 0,58 à 5,25. La prévalence moyenne était de 40 %.

Techniques du scanner

La majorité des études ont utilisé des scanners hélicoïdaux de 3ème génération, le site d'examen était soit abdomino-pelvien soit focalisé. Il existait une variabilité dans les modalités de contraste (sans contraste, oral, intraveineux, rectal ou plusieurs modalités associées).

Le critère de positivité au scanner, adopté dans les études, correspondait à un diamètre de l'appendice supérieur à 6 mm, d'autres critères tels que l'inflammation péri-appendiculaire, l'appendicolithe ou l'épaisseur de la paroi caecale ont également été utilisés.

Analyse de l'hétérogénéité

Il n'a pas été démontré d'hétérogénéité entre les études en termes de sensibilité et spécificité (pour la sensibilité $\chi^2=0,32$ ($p=0,57$), spécificité $\chi^2=0,08$ ($p=0,78$) par conséquent, un modèle à effets fixes a été utilisé pour calculer les sensibilités et spécificité dans la méta-analyse.

Résultats de performance diagnostique

Cette étude rapporte une sensibilité de 0,94 IC 95% [0,92-0,95], une spécificité de 0,94 IC 95% [0,94-0,96] et des rapports de vraisemblance¹⁶ positifs et négatifs respectivement de 15,66 et 0,06 (cf. Tableau 11).

Cette étude montre une performance diagnostique du scanner relativement élevée malgré la variabilité des modalités de contraste utilisées dans les études. Cependant, les biais méthodologiques décrits précédemment (cf. 5.3.1) sont susceptibles d'induire une surestimation des effets, le niveau d'évidence de cette étude en est par conséquent amoindri.

Méta-analyse de Hlibczuk et al. (60)

L'objectif principal de cette revue systématique était d'évaluer la performance diagnostique du scanner réalisé sans produit de contraste chez le sujet adulte admis aux urgences pour suspicion d'appendicite (60). Le test de référence correspondait soit à l'examen macroscopique par le chirurgien soit l'examen anatomopathologique pour les patients opérés soit le suivi clinique d'au moins deux semaines pour les patients non opérés.

Sept études prospectives ont été analysées représentant un total 1060 patients.

Caractéristiques des patients

Dans cette méta-analyse, la population était représentée par des patients admis aux urgences pour douleur abdominale aiguë et suspicion d'appendicite, l'âge des patients dans les études originales était présenté sous formes d'intervalles, la borne inférieure variait de 16 à 23 ans et la borne supérieure variait de 65 à 86 ans. L'âge moyen des patients dans les études ainsi que le sex-ratio (femmes/hommes) n'ont pas été renseignés.

La prévalence médiane était de 39,3% elle variait entre 20,1 % et 84,5% selon les études.

Analyse de l'hétérogénéité

L'hétérogénéité dans cette étude a été testée au moyen de courbes SROC, modèle de régression dans lequel, le log (Odds Ratio) est modélisé (de manière linéaire) par une fonction de la sensibilité et de la spécificité reflétant la fréquence avec laquelle le test diagnostique est positif.

¹⁶ Calculés à partir des valeurs compilées de sensibilité et de spécificité

L'analyse de la sensibilité a été réalisée avec un modèle à effets mixtes qui permet l'estimation des valeurs moyennes de sensibilité et spécificité et permet l'ajustement des sources de biais et de la variabilité inter-études¹⁷.

L'analyse des courbes SROC, n'a révélé aucun effet de seuil significatif $\square = -0,3$ IC 95 % [-1,9 à 1,3]

Techniques de scanner

Dans cette méta-analyse, seules les études qui utilisaient une technologie de scanner hélicoïdal multi-barrettes ont été analysées étant considéré par les auteurs que la technique mono-barrette n'était plus utilisée aux USA. La région de l'exploration (focale, abdomino-pelvienne ou pelvienne), le critère de positivité au scanner adopté dans les études de même que l'expérience du radiologue n'ont pas été renseignés.

Résultats de performance diagnostique

Les valeurs de sensibilité et de spécificité estimés selon un modèle à effets aléatoires sont respectivement de 92,7% IC 95% [89,5-95%] et de 96,1% IC 95% [94,2-97,5]. Les valeurs de rapports de vraisemblances calculées à partir des valeurs compilées de sensibilité et de spécificité sont de 24 pour le RVP et de 0,08 pour le RVN (cf. Tableau 11).

Les principales limites méthodologiques

Les principales limites méthodologiques relevées dans cette étude sont l'hétérogénéité des patients des études. Cette hétérogénéité transparait dans la variabilité du taux de prévalence entre les différentes études avec deux études avec des prévalences particulièrement élevées (77,6 et 84,5 %).

Les études originales (sauf une étude chez qui la laparoscopie diagnostique a été réalisée chez tous les patients) souffrent d'un biais de vérification différentiel, d'un biais d'absence d'insu pour le test de référence (anatomopathologie). Le taux de résultats d'examen de scanner non interprétables n'ont été rapportés que dans 2 études, les autres études les considéraient comme d'emblée négatifs. Les biais de publication n'ont pas été vérifiés ce qui pourrait, si ces biais existaient, induire une surestimation des résultats.

Cette étude montre une bonne performance diagnostique du scanner réalisé sans produit de contraste. Cependant, les limites méthodologiques décrites ci-dessus sont susceptibles d'induire une surestimation des effets.

► Conclusion

A l'analyse de la littérature, il apparait qu'en règle générale, les études originales présentent des biais majeurs susceptibles d'induire une surestimation des valeurs de la performance diagnostique du scanner. Une grande hétérogénéité a été relevée dans les études en ce qui concerne la présentation clinique, le degré de sévérité du tableau clinique, la variabilité des techniques de scanner, les modalités de contraste et les sites d'examen.

En dépit de cette hétérogénéité, les données restent significativement homogènes et montrent une bonne performance diagnostique aussi bien pour confirmer le diagnostic d'appendicite que pour l'infirmier.

¹⁷ Cette technique est devenue la méthode de référence pour réaliser les méta-analyses d'études diagnostiques

5.3.3 Comparaison directe de la performance diagnostique du scanner et de l'échographie

Parmi les études analysées en 5.3.1 et 5.3.2, seule la méta-analyse de van Randen et al. (58) autorisait une comparaison des deux techniques d'imagerie car elles comparaient directement les deux modalités chez le même patient. La méta-analyse de Terasawa et al. (48) présentait également les résultats d'une méta-analyse réalisée sur un sous-groupe d'études (4 études) qui évaluaient également les deux modalités sur le même patient dont 3 sont analysées dans la méta-analyse de van Randen et al..

Au regard des résultats présentés sur le tableau 13, le scanner montre une performance diagnostique supérieure à celle de l'échographie.

Tableau 13. Comparaison directe des résultats de performance diagnostique du scanner et de l'échographie.

	Performance diagnostique du scanner		Performance diagnostique de l'échographie	
	RVP [IC 95%]	RVN [IC 95%]	RVP [IC 95%]	RVN [IC 95%]
van Randen et al., 2008 (58)	9,3 [6,86-12,6]	0,1 [0,06-0,17]	4,5 [3,03-6,68]	0,27 [0,17-0,43]
Terasawa et al. 2004 (48)	12,2 [7,1-21,2]	0,05 [0,02-0,12]	3,4 [2,4-4,8]	0,33 [0,19-0,55]

5.3.4 Les recommandations

► Recommandations de l'ANDEM 1996 (38)

- En 1996, l'Andem avait conclu pour l'échographie, que :
- L'étude échographique de l'appendicite peut être utile lorsque le diagnostic clinique est difficile ou douteux ;
- Une échographie appendiculaire jugée normale n'élimine pas le diagnostic d'appendicite ;
- L'échographie abdomino-pelvienne a un rôle dans le diagnostic différentiel de l'appendicite et dans celui des abcès appendiculaires.

► Les recommandations de l'ACEP 2010

La recommandation américaine de l'ACEP (53) s'adresse à une population de patients adultes se présentant avec une douleur abdominale aiguë non traumatique avec suspicion d'appendicite. L'objectif et la méthode d'élaboration de ces recommandations sont décrits en §5.2.1

En ce qui concerne l'imagerie, l'une des questions à laquelle visait à répondre ce travail, est la suivante :

« Chez les patients adultes avec suspicion d'appendicite chez qui un scanner doit être réalisé, quel est le rôle du produit de contraste ? »

L'ACEP s'est appuyée sur les résultats de 3 études prospectives comparatives (62,66,77) pour répondre à cette question. Chacune des études a été construite selon un schéma comparatif différent.

L'étude de Hershko et al. (62), a été réalisée sur 232 patients, elle comparait selon un schéma contrôlé randomisé 3 groupes de patients chez qui a été réalisé un scanner avec 3 modalités de contraste différentes : 70 patients sans contraste, 78 patients avec un produit de contraste rectal, 84 patients avec contraste intraveineux et oral.

Cette étude montre que l'association d'un contraste oral et intraveineux (O+IV) a une sensibilité et un rapport de vraisemblance plus élevés que les 2 autres modalités avec :

- Une sensibilité de 100% pour le contraste (O+IV) vs 93% pour le contraste rectal ($p < 0,05$) vs 90% pour le scanner sans contraste (p non significatif) ;
- Un RVN respectivement de 0,0 pour le contraste (O+IV) vs 0,07 pour le contraste rectal ($p < 0,05$) vs 0,12 pour le scanner sans contraste (p non significatif).

Le taux de résultats équivoques pour la procédure sans produit de contraste est estimé à 20 %.

Cette étude est classée en niveau d'évidence 2.

L'étude de Jacobs et al. (77), a été réalisée sur 210 patients selon un plan d'étude croisé (case-crossover-design) qui comparait chez le même patient une modalité de scanner focal avec un produit de contraste oral versus un scanner abdomino-pelvien avec produits de contraste (O+IV). Les auteurs attribuent une meilleure performance diagnostique à cette dernière modalité avec une Se= 91% vs 76% pour la modalité de scanner focal avec produit de contraste oral et un RVN de 0,1 vs 0,26. Le degré de significativité n'est pas renseigné.

Cette étude a un niveau d'évidence de 1.

L'étude de Mitall et al. (66) a été réalisée sur 91 patients, elle comparait selon un schéma contrôlé randomisé une modalité de scanner abdomino-pelvien avec un triple contraste (oral, IV et rectal) à un scanner focal avec uniquement un produit de contraste rectal. Les résultats de cette étude montrent que le scanner avec une triple modalité de contraste a une Se de 97%, une Sp de 86%, un RVP de 6.9 et un RVN de 0,04 alors que le scanner focal avec contraste rectal présente une Se de 88%, une Sp de 100%, un RVN de 0,12 (RVP non défini). Cependant les auteurs n'ont pas rapporté le degré de significativité des résultats. Cette étude souffre, par ailleurs, de nombreuses limites méthodologiques notamment la différence de caractéristiques de population entre les deux groupes en termes : d'âge, de sex-ratio, de degré de sévérité (taux d'appendicites perforées) et de délai avant la réalisation d'examen d'imagerie.

Cette étude a un niveau d'évidence de 3.

En conclusion, concernant le rôle du produit de contraste chez les patients qui doivent réaliser un scanner pour suspicion d'appendicite, l'**ACEP** recommande :

« Chez les adultes les patients qui réalisent un scanner pour suspicion d'appendicite, « effectuez » un scanner abdomino-pelvien avec ou sans produit de contraste (par voie intraveineuse (IV), par voie orale ou rectale) l'ajout d'un produit de contraste par voie IV et par voie orale peut augmenter la sensibilité de scanner pour le diagnostic de l'appendicite ». **Niveau B**

► Les recommandations de l'American College of Radiology

ACR appropriateness Criteria®: Right Lower Quadrant Pain- Suspected Appendicitis (36)

Ces recommandations de bonne pratique ont été élaborées selon les critères de l'agence américaine de recherche sur la qualité des soins de santé : Agency for Healthcare Research and Quality (**AHRQ**). Elles reposent sur une méthode de consensus formalisé (adaptation de la méthode DELPHI¹⁸) qui s'appuie sur une recherche exhaustive de la littérature et sur la cotation des conclusions issues de l'analyse de la littérature par un groupe d'experts ;

La population cible correspond aux patients avec une douleur abdominale en fosse iliaque droite et suspicion d'appendicite. Conformément au champ défini pour ce travail, seules les recommandations concernant le patient adulte ont été considérées.

Ce travail de l'ACR visait à comparer les performances diagnostiques du scanner et de l'échographie ainsi que la performance diagnostique des différentes modalités de contraste pour le scanner.

Pour le premier point, les auteurs se sont appuyés principalement sur la méta-analyse de van Randen et al. (58) (cf. 5.3.3), seule méta-analyse de la littérature à avoir réalisé une comparaison directe de la performance diagnostique de l'échographie et du scanner sur le même patient. Une meilleure précision diagnostique a été retrouvée pour le scanner (cf. tableau 12) avec une cohérence des résultats entre les différentes études alors que pour l'échographie, la grande hétérogénéité qui a été retrouvée entre les études suggèrerait un caractère très opérateur dépendant de cette technique.

Pour le deuxième point qui concerne la performance diagnostique des différentes modalités de contraste, le travail de l'ACR s'est appuyé principalement sur les résultats de performance diagnostiques l'étude de Hershko et al. (62) décrite plus haut (cf. recommandations de l'ACEP) qui ont montré qu'une meilleure performance diagnostique était obtenue en cas d'utilisation d'un produit de contraste par injection intraveineuse. L'utilisation simultanée ou pas d'un produit de contraste par voie orale ne semblait pas modifier les valeurs de sensibilité et spécificité.

En cas de contraindication à l'utilisation d'un produit de contraste, l'ACR s'est appuyé sur l'étude prospective consécutive (300 patients) de Lane et al (78) (niveau d'évidence 2) pour montrer que le scanner sans contraste restait une modalité « appropriée » avec une précision diagnostique de 96%, une sensibilité de 96% et une spécificité de 99%.

L'ACR a établi une échelle d'évaluation qui permet de classer de 1 à 9 chaque modalité d'imagerie en fonction de son adéquation avec une situation clinique donnée. La signification des scores est la suivante :

- **1, 2,3** habituellement pas approprié ;
- **4, 5,6** peut être approprié ;
- **7, 8,9** habituellement appropriée.

L'ACR distingue deux situations cliniques différentes pour le patient adulte qui présente une douleur en FID et une suspicion d'appendicite. En fonction de chaque situation, différents scores sont attribués aux modalités d'imagerie.

Première situation : Douleur en FID, suspicion d'appendicite

Présence simultanée de fièvre, d'hyperleucocytose et d'une présentation clinique de l'appendicite dite « classique »

Seules les techniques qui « sont habituellement appropriées ou peuvent être appropriées » sont résumées dans le tableau ci-dessous :

¹⁸ La méthode Delphi a pour but de mettre en évidence des convergences d'opinion et de dégager certains consensus sur des sujets précis, grâce à l'interrogation d'experts, à l'aide de questionnaires successifs.

Modalité d'imagerie	score	Commentaires
Scanner abdomino-pelvien avec contraste	8	L'utilisation d'un contraste oral ou rectal dépend de la préférence de l'institution
Scanner abdomino-pelvien sans contraste	7	L'utilisation d'un contraste oral ou rectal dépend de la préférence de l'institution
Echographie abdominale en FID	6	Avec une compression graduée
Echographie pelvienne	5	
Radiographie de l'abdomen	5	Peut être utile pour éliminer l'air libre ou l'obstruction.

Deuxième situation : Douleur en FID, suspicion d'appendicite

Présence simultanée de fièvre, d'hyperleucocytose, d'une appendicite possible mais avec une présentation clinique dite « atypique »

Modalité d'imagerie	score	Commentaires
Scanner abdomino-pelvien avec contraste	8	L'utilisation d'un contraste oral ou rectal dépend de la préférence de l'institution
Scanner abdomino-pelvien sans contraste	6	L'utilisation d'un contraste oral ou rectal dépend de la préférence de l'institution
Echographie abdominale en FID	6	Avec une compression graduée
Echographie pelvienne	6	
Radiographie de l'abdomen	6	Peut être utile pour éliminer l'air libre ou l'obstruction.

Quelle que soit la situation clinique, l'ACR considère que le scanner avec ou sans contraste est la modalité d'imagerie qui est « habituellement appropriée ». Cependant la classification des méthodes d'imagerie en fonction de leur pouvoir irradiant chez l'adulte, classe le scanner en première position avec une dose de 30 à 100 mSv contre 1 à 10 mSv pour la radiographie et l'absence totale d'irradiation pour l'échographie.

En résumé ce travail de l'ACR conclut que :

L'appendicite peut être diagnostiquée cliniquement, mais l'imagerie augmente la sensibilité et la spécificité du diagnostic.

En général, le scanner est la modalité d'imagerie la plus précise pour évaluer une suspicion d'appendicite et les étiologies possibles de la douleur en fosse iliaque droite.

► Conclusion des recommandations

De l'analyse de ces recommandations, il ressort les éléments suivants :

- l'appendicite peut être diagnostiquée cliniquement, mais l'imagerie augmente la précision diagnostique ;
- l'échographie n'a de valeur que positive, une échographie appendiculaire jugée normale n'élimine pas le diagnostic d'appendicite ;
- le scanner est plus précis quelle que soit sa modalité de réalisation (avec ou sans produit de contraste), il présente néanmoins un plus grand risque d'irradiation.

5.4 La coelioscopie exploratrice

Pour l'évaluation de la coelioscopie exploratrice, après sélection de la littérature identifiée, les documents suivants ont été retenus :

- La conférence de consensus de l'European Association for Endoscopic Surgery (EAES) (79) ;
- Les recommandations de la Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) (80) ;
- Une revue Cochrane (81) ;
- Une conférence de consensus italienne de l'Associazione dei Chirurghi Ospedalieri Italiani (ACOI) (82).
- Parmi les études exclues de l'analyse, trois études prospectives randomisées (83-85) qui ont déjà été analysées dans la revue Cochrane et une revue systématique (86) dont la période de recherche documentaire était antérieure à celle de la revue Cochrane.

5.4.1 EAES Consensus Statement 2006

En 2006 l'EAES a organisé une conférence de consensus européenne dont les recommandations ont été publiées par Sauerland et al. (79). L'objectif de ce travail était de définir, pour les pathologies les plus fréquentes, les sous groupes de patients pouvant bénéficier de la coelioscopie. Il visait principalement à comparer l'efficacité de la chirurgie coelioscopique par rapport à la laparotomie, les aspects diagnostiques de la coelioscopie ont été évoqués secondairement. La population cible concernée consistait en des patients adultes présentant une douleur abdominale aiguë évoluant depuis moins de 7 jours.

Dans ce travail, la sélection des articles a privilégié les essais contrôlés randomisés de niveau 1 ou non randomisés de niveau 2 et le cas échéant a eu recours aux études de moindre niveau de preuves. Sur la base de la revue de la littérature, les textes des recommandations ont été élaborés et validés selon la méthode DELPHI¹⁹.

Concernant l'appendicite aiguë, le texte des recommandations s'est appuyé essentiellement sur une précédente revue systématique du même auteur publiée en 2004 (87) et sur un essai randomisé publié en 2005 (88). Ce travail étant principalement orienté vers la chirurgie coelioscopique, seules quelques études (6 sur 56 études) ont rapporté les résultats de l'exploration coelioscopique à visée diagnostique qui ont été utilisés pour décider d'une chirurgie ultérieure.

Les auteurs rapportent, sans étayer les affirmations par des valeurs quantitatives, une prédominance de femmes jeunes dans les études et un taux important de diminution du taux d'appendicectomies négatives chez cette même population.

Ils rajoutent par ailleurs que les avantages de la coelioscopie exploratrice chez les hommes et les enfants semblent être moindres du fait d'un diagnostic d'appendicite plus facile à établir dans ces populations.

¹⁹ La méthode Delphi a pour but de mettre en évidence des convergences d'opinion et de dégager certains consensus sur des sujets précis, grâce à l'interrogation d'experts, à l'aide de questionnaires successifs.

Pour la coelioscopie exploratrice de l'appendicite, une seule recommandation est issue de ce travail, elle est formulée comme suit :

« Les patients présentant des symptômes et des éléments de diagnostic suggérant une appendicite aiguë devraient subir une laparoscopie (coelioscopie) diagnostique (grade A) »

Au regard de la population recrutée dans les études, majoritairement des femmes jeunes, une telle formulation qui généralise l'usage de la coelioscopie exploratrice à tous les patients ne correspond pas aux informations analysées. Malgré la gradation attribuée (garde A cf. annexe 8), cette recommandation semble plus relever de l'avis d'experts que reposer sur des données factuelles de la littérature.

Par ailleurs, il apparaît dans la conclusion du résumé de l'article une précision qui nuance les propos de la recommandation en positionnant la coelioscopie exploratrice après les examens diagnostiques conventionnels, cette formulation n'a pas été retrouvée dans le texte intégral,

« En cas de douleur abdominale aiguë stable, il faut épuiser toutes les méthodes diagnostiques non invasives avant d'avoir recours à la laparoscopie. Celle-ci ne doit être préconisée que si les procédures classiques de diagnostic n'ont pas réussi à donner des résultats »

Sur le plan méthodologique, la qualité des études de même que les critères de sélection et d'analyse des articles n'ont pas été décrits, l'argumentaire très peu précisé et les recommandations telles que formulées semblent élargir la pratique à une population générale alors même que les données ont été analysées sur une cible particulière. Par ailleurs, la place de la coelioscopie exploratrice dans la stratégie d'exploration morphologique ne paraît pas claire dans le texte des recommandations.

5.4.2 Recommandations de SAGES 2007 (Society of American Gastrointestinal and Endoscopic Surgeons)

Les recommandations publiées en 2007 par la SAGES sont une mise à jour de lignes directrices antérieures publiées en 2002 (80). Elles avaient pour objectif l'élaboration d'une série de recommandations pour aider les chirurgiens à prendre une décision appropriée quant à l'utilisation de la laparoscopie à visée diagnostique dans des situations de douleur abdominale aiguë évoluant depuis moins de 7 jours.

Ces recommandations sont issues d'une revue systématique de la littérature (13 études de niveau de preuves variable (1 à 3) dont 2 essais contrôlés randomisés) et de l'avis d'experts. Les forces et les faiblesses des études ont été évaluées par le comité d'experts qui a octroyé un grade aux recommandations en fonction du niveau de preuves des études qui supportaient les affirmations. Pour les points pour lesquels la littérature fait défaut, il a été fait recours uniquement à l'avis d'experts.

Sur les 13 études analysées, l'efficacité diagnostique de la laparoscopie variait de 70 à 99%.

Sur le plan méthodologique, ce travail présente une validité relative, il faut noter que seuls les travaux publiés en langue anglaise ont été recherchés, les critères d'inclusion et d'exclusion des études n'ont pas été précisés de même que le mode de fonctionnement des groupes de travail et des comités d'experts. Sur les données recueillies, les auteurs rapportent également une difficulté à compiler les résultats du fait d'une hétérogénéité des patients et de l'expérience des opérateurs.

Les conclusions émises à l'issue de ce travail ne sont par ailleurs pas spécifiques au diagnostic d'appendicite mais regroupent de manière générale les situations de douleur abdominale aiguë. Dans cette situation la SAGES a émis les recommandations ci-dessous, deux d'entre elles au vu du niveau de gradation attribué (C) reposent uniquement sur avis d'experts.

- La coelioscopie diagnostique est une technique faisable et peut être appliquée en toute sécurité chez les patients avec une douleur abdominale aiguë non spécifique, sélectionnés de manière appropriée (Grade B) ;

- La procédure doit être évitée chez les patients avec une instabilité hémodynamique et limitée chez les patients avec une distension abdominale sévère ou avec une indication claire de laparotomie (Grade C) ;
- La procédure ne peut être envisagée que chez les patients dont le diagnostic reste imprécis après examen clinique et examens d'imagerie (Grade C).

Les auteurs précisent qu'à la lumière des données disponibles, une procédure invasive telle la coelioscopie exploratrice ne peut être recommandée avant que toutes les options diagnostiques non invasives n'aient été épuisées.

5.4.3 Revue systématique Cochrane 2010

Une revue Cochrane a été publiée en 2010 (81) sur le thème « La chirurgie laparoscopique versus la chirurgie conventionnelle ouverte dans la suspicion d'appendicite », Il s'agit d'une mise à jour publiée en 2006 par les mêmes auteurs et dont l'objectif principal était de comparer les effets de diagnostic et de thérapeutique de la laparoscopie et de la chirurgie conventionnelle ouverte.

En raison de l'orientation du présent travail, seul l'aspect diagnostic de cette revue a été appréhendé. Pour cet item, les auteurs ont sélectionné 14 études comparatives randomisées, publiées sur la période allant de 2003 au 15 avril 2010, portant sur des patients présentant des symptômes et des signes d'appendicite aiguë et rapportant un taux d'appendicites avec signes histologiques d'inflammation de plus de 50%.

Les procédures comparées étaient les suivantes :

- Appendicectomie conventionnelle comparée à la coelioscopie exploratrice suivie par l'appendicectomie coelioscopique ou conventionnelle si nécessaire ;
- Coelioscopie exploratrice (suivie par l'appendicectomie ouverte si nécessaire) comparée à une appendicectomie ouverte immédiate seule.

Les critères de jugement pour la coelioscopie exploratrice étaient :

- le taux d'appendicectomies négatives ;
- le taux de patients dont le diagnostic n'est pas établi (non conclusif).

Les études ont été séparées selon qu'elles s'adressaient à des patients « non sélectionnés » ou à des femmes jeunes en âge de procréer.

Dans les essais sur des patients non sélectionnés (8 essais, 667 patients) une diminution du taux d'appendicites négatives a été constatée avec un RR 0,37, IC 95% [0,13-1,01] mais le résultat global a été très influencé par la puissance d'un seul essai (282 patients, RR 0,05). Parallèlement le taux de diagnostics non conclusifs a également diminué après coelioscopie exploratrice avec un RR de 0,43 IC 95% [0,17-1,08]. Au regard des intervalles de confiance, ces résultats ne sont statistiquement pas significatifs. Une hétérogénéité importante a par ailleurs été relevée avec un test I² =74% (p<0,053).

Dans les études qui s'adressent aux femmes jeunes les résultats de la coelioscopie exploratrice montrent également une réduction du nombre d'appendicectomies inutiles avec un RR de 0,20 IC 95% [0,11-0,34] de même que le nombre diagnostic non conclusifs (RR 0,27 ; IC 95% [0,17 -0,44]. Ces résultats sont basés sur six essais similaires qui à l'analyse ne révélaient aucune hétérogénéité (I²=0% (p<0,00001).

Les auteurs concluent que dans les deux groupes l'utilisation de la coelioscopie exploratrice réduit le risque des appendicectomies négatives même si l'effet est bien plus prononcé chez les jeunes femmes. Ils considèrent que cette efficacité pourrait être sous-évaluée arguant du fait que dans les études analysées, excepté les situations où la cause de douleur abdominale est identifiée, l'exérèse d'un appendice même macroscopiquement sain est quasi systématique, ce qui augmente le taux d'appendicectomies négatives.

Les conclusions des auteurs sont les suivantes :

« Dans les situations cliniques où l'expertise chirurgicale et l'équipement sont disponibles et abordables, il est recommandé d'utiliser en règle générale la laparoscopie diagnostique et l'appendicectomie laparoscopique chez tous les patients avec une suspicion d'appendicite à moins que la technique ne soit contre-indiquée ou non réalisable. Cette recommandation, toutefois, ne peut s'appliquer aux appendicites perforées ou gangreneuses en raison du risque de survenue d'abcès intra-abdominaux après le geste laparoscopique »

Ils rajoutent par ailleurs que cette recommandation s'adresse plus spécifiquement aux femmes jeunes et aux personnes obèses chez qui les avantages semblent être plus évidents.

Sur le plan méthodologique, la qualité de toutes les études incluses dans cette revue est modérée voire faible. Les critères cliniques pour identifier les patients avec suspicion d'appendicite aiguë sont subjectifs, l'intervalle d'âge, le sex-ratio (plusieurs études ont inclus uniquement une population féminine) et l'expérience des opérateurs sont très variables. Les critères de réalisation d'une appendicectomie après coelioscopie exploratrice sont très variables selon les études.

Une hétérogénéité importante a été constatée notamment pour les études de patients non sélectionnées.

Plusieurs biais ont été relevés dans les études, notamment : une génération de séquence de randomisation inadéquate ou non décrite, une absence de dissimulation de l'affectation des sujets aux groupes d'étude, une absence d'interprétation des résultats en insu, des données de perdus de vue non prises en compte, un biais d'interprétation et de jugement.

Au regard de tous ces éléments, les conclusions des auteurs semblent non conformes aux données recueillies. De plus, au-delà de l'objectif attendu d'une revue systématique, les auteurs émettent des recommandations qui reposent d'avantage sur une conviction personnelle que sur une réalité factuelle qui ressortirait de l'analyse de la littérature.

Hormis le cas des femmes jeunes pour lequel les données semblent favorables, la généralisation de la laparoscopie diagnostique à toutes les populations est, au regard des données non acceptable.

5.4.4 Conférence de consensus de l'association italienne des chirurgiens hospitaliers (ACOI) 2011

Les recommandations issues de la conférence de consensus de l'ACOI (82) ont été élaborées selon les critères de l'American National Institute of Health (NIH) qui préconisent l'utilisation de la méthode de DELPHI²⁰. La stratégie de recherche documentaire a privilégié les essais contrôlés randomisés et les méta-analyses et le cas échéant des études de moindre niveau de preuves. Un groupe de 20 chirurgiens italiens ont été désignés par le comité directeur de l'ACOI pour analyser la littérature et rédiger les propositions de recommandations. Celles-ci ont été discutées lors du congrès nationale de l'ACOI et l'avis de l'audience a été recueilli par vote électronique. Ces recommandations ont été ensuite approuvées par le comité d'experts qui leur a attribué des niveaux de preuve et ont été validées en fin de parcours par deux experts chirurgiens reconnus par leur pairs.

L'objectif de cette conférence, tel que défini par les auteurs, était d'apporter des réponses « basées sur les preuves » à la controverse concernant l'utilisation de la laparoscopie dans le traitement de l'appendicite. Le champ de ce travail n'était pas clairement précisé, plusieurs aspects ont été abordés, ils ont donné lieu à des recommandations concernant : l'algorithme diagnostique, la coelioscopie exploratrice, les indications, la conduite à tenir en cas d'appendice sain, les aspects techniques et la courbe d'apprentissage.

²⁰ La méthode Delphi a pour but de mettre en évidence des convergences d'opinion et de dégager certains consensus sur des sujets précis, grâce à l'interrogation d'experts, à l'aide de questionnaires successifs.

Sur le plan méthodologique, ce travail présente une validité relative, il n'y a pas de définition précise de l'objectif du travail, la population cible est très peu décrite, il s'agit d'une population adulte mais aucune caractéristique clinique permettant de définir la suspicion d'appendicite n'est décrite, seuls les travaux publiés en langue anglaise sont recherchés, les critères d'inclusion et d'exclusion des études ne sont pas précisés, les critères de jugement adoptés dans les études ne sont pas décrits.

Les lignes directrices consensuelles publiées à l'issue de cette conférence ont défini la place de la laparoscopie dans la stratégie diagnostique. Il en est ressorti que :

- le diagnostic de l'appendicite est essentiellement clinique,
- que l'échographie est complémentaire pour les diagnostics différentiels ;
- que le scanner est un examen de second choix en particulier chez les sujets âgés et personnes obèses.

Ainsi, la coelioscopie exploratrice est positionnée en dernier recours quand les autres solutions diagnostiques sont non conclusives.

Deux « affirmations » sur la coelioscopie exploratrice sont ressorties de cette conférence de consensus :

- La coelioscopie exploratrice est recommandée pour toutes les douleurs abdominales aiguës situées dans la fosse iliaque droite, elle est considérée comme un diagnostic « final » et un acte thérapeutique formel. (Niveau A).

Pour cette affirmation, les auteurs s'appuient essentiellement sur la revue Cochrane (81) et sur les recommandations de Sages décrites plus haut.

- Les femmes jeunes en âge de procréer (niveau A) et les groupes à risque de complications majeures (patients âgés et obèses ; Niveau C) devraient bénéficier le plus de cette approche.

► Conclusion

En s'appuyant sur l'analyse des différents documents de recommandations, il apparaît qu'il n'y a pas de données de qualité suffisante pour positionner la coelioscopie exploratrice dans la démarche diagnostique. La coelioscopie exploratrice ne doit pas être réalisée en première intention, elle peut être proposée lorsque l'imagerie (échographie puis scanner) n'est pas contributive, chez la femme jeune.

5.5 Le traitement antibiotique versus l'appendicectomie.

Pour la comparaison du traitement conservateur et de l'appendicectomie, seuls ont été retenus les méta-analyses d'essais contrôlés randomisés dont la période de recherche de la littérature allait jusqu'en 2011. Deux documents ont été identifiés,

- une revue systématique Cochrane (89) ;
- une méta-analyse publiée en 2012 par Varadhan et al. (90).

En raison de la similarité des critères de sélection de la littérature, ces deux documents de synthèse ont sélectionné deux études en commun.

5.5.1 Méta-analyse de Varadhan et al. (90)

Il s'agit d'une réactualisation d'un précédent travail publié en 2010 par le même auteur (14) qui a inclut la récente étude française publiée en 2011 (16).

L'objectif de cette méta-analyse était de comparer, pour les appendicites non compliquées, la sécurité et l'efficacité du traitement antibiotique comme traitement de première intention à l'appendicectomie. Le critère de jugement principal étant les complications,

Selon les deux groupes de patients, la définition des complications est différente :

- pour le groupe antibiotiques, les complications correspondaient à une appendicite perforée ou gangréneuse et en cas d'appendicectomie secondaire à une péritonite et une infection de la plaie ;
- pour le groupe appendicectomie, les complications correspondaient à une péritonite et une infection de la plaie.

Pour la comparaison de l'efficacité des deux traitements, cette méta-analyse a adopté un critère composite qui comprenait : le nombre d'appendicites compliquées, la durée du séjour et le nombre de ré-hospitalisations sur une période de suivi d'un an.

Quatre essais contrôlés randomisés ont été inclus dans cette méta-analyse ce qui représentait un total de 900 patients, dont 470 ont été randomisés dans le bras traitement antibiotique et 430 dans le bras appendicectomie.

► **Caractéristiques des patients**

La moyenne d'âge des patients variait dans 3 études pour le groupe appendicectomie entre 35 et 38 ans et pour le groupe antibiothérapie entre 27 et 38 ans. Dans la quatrième étude l'âge des patients était présenté sous forme de bornes comprises entre 18 et 50 ans. Le sex-ratio (femmes/hommes) n'était pas renseigné et dans une étude la population incluse était exclusivement masculine.

La population des études était représentée par des patients avec une « appendicite non compliquée » sélectionnés sur la base de paramètres diagnostiques cliniques et biologiques. Dans deux études l'imagerie était réalisée systématiquement chez tous les patients, dans une troisième étude le diagnostic reposait uniquement sur les signes cliniques et dans une quatrième étude le diagnostic était clinique mais parfois il était fait appel à l'imagerie (échographie ou scanner).

► **Traitements antibiotiques**

Les traitements ainsi que les schémas posologiques sont variables dans les études, dans deux études américaines, il a été utilisé une association de céfotaxime et tinidazole en intraveineux à l'hôpital, elle a été complétée par association d'ofloxacine et tinidazole par voie orale à la sortie. Une troisième étude américaine a utilisé une association de céfotaxime et métronidazole en intraveineux complétée par une association ciprofloxacine et metronidazole en prise orale à la sortie de l'hôpital.

Dans l'étude française de Vons et al. (16), il a été utilisé une association d'amoxicilline et d'acide clavulanique.

► **Analyse de l'hétérogénéité**

L'hétérogénéité dans cette étude a été testée au moyen du test I^2 de Higgins qui calcule le pourcentage de variation entre les études lié à une hétérogénéité et non au hasard.

Pour les complications aucune hétérogénéité n'a été retrouvée ($I^2 = 0$), le rapport de risques (RR) a été calculé selon un modèle à effets fixes (IC 95%). Pour l'efficacité, une hétérogénéité a été retrouvée (sur l'ensemble des études $I^2 = 82\%$ et en retirant les résultats en « cross-over $I^2 = 74\%$), le rapport de risques a été calculé selon un modèle à effets aléatoires (IC 95%) .

Les résultats montrent une réduction du risque de complications de 31% par le traitement antibiotique par rapport à l'appendicectomie soit un rapport de risques de 0,69 IC 95% [0,54-0,89] $p=0,004$. Une deuxième analyse des données a été effectuée en retirant les résultats en « cross-over » (patients sous antibiotiques qui ont subi une appendicectomie secondaire). Les données ainsi analysés ont montré une diminution de risque de complications de l'ordre de 39% soit un $RR=0,61$ IC 95% [0,40-0,92] $p=0,02$.

Pour l'efficacité du traitement, les auteurs rapportent un risque d'appendicites compliquées similaire dans les deux groupes avec un RR= 0,58 IC 95% [0,18-1,90]. La comparaison de la durée de séjours ne montre pas non plus de différence entre les deux groupes avec une valeur de différence moyenne de 0,34 IC95% [-0,19 à 0,87]. Le taux de rechutes à un an est de 20% chez les patients traités par antibiotiques dont 19% présentent une appendicite compliquée.

► Les principales limites méthodologiques

Les études originales sont toutes de qualité faible à moyenne, elles souffrent principalement de biais d'attrition et de biais de sélection.

Une étude a inclus exclusivement des hommes âgés entre 18 et 50 ans, une autre étude a présenté des résultats en « cross-over » avec superposition des résultats du groupe antibiotiques avec le groupe appendicectomie.

Un biais de publication a été mis en évidence par le « funnel plot » du fait du faible nombre d'études incluses (quatre), mais les auteurs ont précisé que la recherche des données non publiées n'a pas révélé de données négatives.

Plusieurs facteurs confondants peuvent également être soulignés notamment l'absence de définition claire d'une « appendicite non compliquée », les critères de sélection d'une appendicite compliquée ne sont pas uniformes entre les études.

L'étude française qui a été incluse dans la méta-analyse (16) utilise des traitements antibiotiques différents (molécules, voie d'administration et durée).

Par ailleurs, la pertinence clinique du critère de jugement peut être discutable, la comparaison en termes de complications est difficile du fait de leur spécificité en fonction du traitement. La rechute est propre aux antibiotiques de la même manière que l'infection de la plaie est propre à la chirurgie.

De plus en raison du biais d'attrition dans les études, il est difficile de recueillir le nombre exact d'appendicectomies subséquentes (rechutes) à un traitement antibiotique.

En conclusion, Cette étude montre une efficacité équivalente entre les deux traitements et une diminution du risque de complications avec l'antibiothérapie. Il faut rappeler cependant que les données de suivi, sur un an, montrent un taux de rechutes de l'ordre de 20% et qu'un suivi plus long pourrait peut être révéler un taux encore plus important. Sur le plan méthodologique, de nombreuses limites confèrent à cette méta-analyse une évidence relative.

En définitive, les résultats semblent prometteurs mais mériteraient d'être confirmés dans le cadre d'études contrôlées randomisées multicentriques avec un délai de suivi plus long.

5.5.2 Revue systématique Cochrane (89)

L'objectif de cette revue systématique était de répondre à la question suivante « est-ce que le traitement antibiotique est aussi efficace que l'appendicectomie (laparoscopique ou laparotomique) chez les patients avec une appendicite aiguë pour permettre la guérison en deux semaines sans complications majeures (y compris la rechute) sur une période de suivi d'un an ? »

Seuls les essais randomisés ou quasi randomisés comparant l'efficacité des deux traitements en termes de supériorité ou non infériorité, chez des patients avec suspicion d'appendicite, ont été inclus.

L'analyse de non infériorité a recherché si l'antibiothérapie était « aussi efficace ou un peu moins efficace » que l'appendicectomie avec un seuil d'infériorité de 20%. Ce seuil qui semble élevé a été consentie au regard des autres avantages de l'antibiothérapie qui sont l'acceptabilité par le patient et le coût.

Le critère de jugement principal est la guérison au bout de deux semaines, définie comme la disparition des signes et symptômes (douleur abdominale, fièvre, marqueurs de l'inflammation) sans complications majeures (y compris la rechute) sur une période de suivi d'un an.

Cinq essais contrôlés randomisés ont été inclus dans cette méta-analyse ce qui représentait un total de 901 patients.

► **Caractéristiques des patients**

L'âge moyen des patients était de 28.2 ans [13-75], le pourcentage d'hommes dans la méta-analyse était de 73,8%. Une étude a inclus une population exclusivement masculine (la même étude que celle analysé dans la méta-analyse de Varadhan et al.).

Comme dans la méta-analyse de Varadhan et al. (90) tous les patients ont été inclus pour suspicion d'appendicite aiguë diagnostiquée sur la base d'éléments cliniques et biologiques (leucocytes et CRP), une imagerie complémentaire a été réalisée uniquement dans 4 études sur cinq : une étude a réalisé un scanner systématique chez tous les patients, deux autres l'échographie chez tous les patients et la quatrième étude a réalisé l'échographie ou le scanner uniquement chez les patients du groupe antibiotiques après randomisation.

► **Traitements antibiotiques**

Dans les études américaines, deux études ont utilisé une association de céfotaxime et tinidazole en intraveineux à l'hôpital, complétée par association d'ofloxacine et tinidazole par voie orale à la sortie. Une troisième étude a utilisé une association de ciprofloxacine et metronidazole en intraveineux à l'hôpital, complétée par association ciprofloxacine et tinidazole par voie orale à la sortie. Une quatrième étude a utilisé une association d'ampicilline et gentamicine et metronidazole en intraveineux à l'hôpital, complétée par la même association d'antibiotiques par voie orale à la sortie.

Dans l'étude française de Vons et al. (16), il a été utilisé une association d'amoxicilline et d'acide clavulanique.

5.5.3 Analyse de l'hétérogénéité

L'évaluation du degré d'hétérogénéité a été réalisée au moyen du test Q de Cochrane et de l'indice I². Quand I² < 30 un modèle effet fixes a été utilisé, quand I² < 50% un modèle à effets aléatoires a été utilisé et en cas d'hétérogénéité importante I² < 75%, aucune compilation des données n'était autorisée.

Concernant l'efficacité du traitement, dans le groupe appendicectomie, 97,4% IC 95% [94,4-98,8] sont guéris au bout de 2 semaines sans complications majeures constatées sur une période de suivi d'un an contre 73,4% 95% [62,7- 81,9] dans le groupe antibiotiques. L'analyse de non infériorité à un seuil de 20% de l'IC 95% montre que la borne inférieure du groupe antibiotiques est de 15,2% inférieur au seuil de 20% et la borne supérieure se superpose au seuil 20% de l'appendicectomie, les résultats en sont par conséquent non concluants.

► **Les principales limites méthodologiques**

Les mêmes limites méthodologiques que celles citées plus haut dans la méta-analyse de Varadhan sont retrouvées, avec principalement une qualité faible à moyenne des études originales. Le sex-ratio est fortement déséquilibré avec près de 74% de population masculine, ce qui représente un biais important du fait d'un plus grand taux d'appendicites négatives chez la femme. Les méthodes diagnostiques sont très différentes entre les études et parfois entre le groupe antibiotique et le groupe appendicectomie. Les méthodes de randomisation sont très peu décrites dans les études originales.

Du fait de la faible qualité des études originales et du faible effectif qui n'autorise pas l'analyse en sous groupe, les résultats de cette méta-analyse ne sont pas concluants. Des études randomisées

de bonne qualité doivent être réalisées en sélectionnant les groupes de patients qui peuvent bénéficier de ce traitement et en augmentant la durée de suivi.

5.5.4 Etudes à venir ou actuellement en cours

Une étude randomisée française est en cours de préparation par l'équipe de Corinne Vons, elle portera notamment sur le sous-groupe de patients avec appendicite aiguë sans stercolithe ayant démontrée une efficacité identique à celle de l'appendicectomie dans la précédente étude randomisée multicentrique (16), le traitement antibiotique devra être amélioré notamment en excluant l'association amoxicilline-acide clavulanique utilisée en 2002 pour la première étude.

Par ailleurs, la recherche d'études cliniques actuellement en cours sur le site de clitrial.gov a retrouvé deux études randomisées :

L'étude « APPAC » NCT01022567 : étude finlandaise sponsorisée par l'hôpital universitaire de Turku dont l'objectif est de comparer l'efficacité de l'appendicectomie vs une antibiothérapie (Ertapenem) chez des patients adultes âgés de 18 à 60 ans dont le diagnostic d'appendicite non compliqué est confirmé par scanner. Le critère de jugement principal étant la résolution de l'appendicite.

L'étude « ASAA » NCT 01421901 : étude italienne sponsorisée par l'hôpital Riuniti di Bergamo dont l'objectif est de tester la non infériorité de l'antibiothérapie par Ertapenem à l'appendicectomie chez des patients adultes âgés de 18 à 65 ans qui ont une probabilité élevée d'appendicite. Le critère de jugement principal étant le taux de patients dont les symptômes ont disparu au bout de 2 semaines (pas de douleur, pas de fièvre, leucocytose $\leq 10.10^9$ et CRP ≤ 1).

► Conclusion

Au regard des données publiées disponibles qui présentent des faiblesses méthodologiques, le traitement antibiotique ne peut actuellement être considéré validé. Des études contrôlées randomisées sont nécessaires pour démontrer l'efficacité des antibiotiques notamment à long terme et pour identifier les sous-groupes de patients pouvant bénéficier de ce traitement.

6. Avis des professionnels

6.1 Concept de la pertinence

Lors des réunions avec les professionnels, il a été rappelé que la qualité des soins étant au centre des missions de la HAS et le concept de pertinence a été défini par les membres du groupe d'orientation (GO) comme : « le juste soin au bon patient, au bon moment, au coût acceptable ».

Les experts ont tenu à souligner que l'enjeu qui sous-tend ce travail est l'adoption d'une attitude diagnostique appropriée qui permette aussi bien de diminuer le risque des appendicites perforées que de réduire le nombre d'appendicectomies inappropriées.

Les membres du GO ont souligné l'importance de la pertinence de l'acte chirurgical en rappelant qu'une appendicectomie inutile n'est pas sans risque. L'apparition fréquente de brides intestinales constitue la première cause d'occlusion post-opératoire (2 à 3%) et représente 40% des causes d'occlusions intestinales aiguës.

Par ailleurs, dans l'évaluation a posteriori de la pertinence d'une appendicectomie, la prise en compte de l'histoire naturelle de l'appendicite est primordiale selon les experts. Les formes d'appendicite sont très variables : elles peuvent être graves d'emblée, moins graves ou résolutive.

Le rôle de l'anatomo-cytopathologiste est prépondérant pour définir la signification histopathologique de chaque forme et confirmer ou infirmer une appendicite.

Selon les experts du groupe d'orientation, les questions suivantes méritent d'être considérées :

- L'analyse de la concordance entre l'aspect histologique de la pièce d'exérèse et le tableau clinique faisant suspecter une appendicite ;
- La caractérisation histopathologique de la pièce d'exérèse.

Les experts ont souligné, à ce propos, l'intérêt de disposer d'une définition histopathologique ad-hoc d'une appendicite aiguë avec un seuil à partir duquel on peut considérer qu'un appendice est pathologique. Les experts rapportent que parfois le seul traumatisme chirurgical peut induire un afflux de polynucléaires et suggérer une inflammation. Vraisemblablement, quand l'appendicite n'a pas une présentation macroscopique évidente (appendicites suppurées...), une variabilité inter-observateur parmi les pathologistes pourrait exister dans le diagnostic d'une appendicite débutante.

Enfin, il a été rappelé, que l'objectif de ce travail était d'élaborer des référentiels pour guider la décision clinique. Il a été souligné la difficulté d'élaborer des propositions claires dans un contexte où il existe une hétérogénéité des données recueillies, un faible niveau de preuves, un manque d'engouement pour cette « vieille affection », une absence de données scientifiques nouvelles et un indicateur de référence discuté (analyse histopathologique).

6.2 Evolution du nombre d'appendicectomies et variabilité interrégionale

Interrogés sur l'évolution du nombre d'appendicectomies sur les dernières années, les professionnels de santé considèrent que la diminution du nombre global d'appendicectomies pourrait s'expliquer par :

- l'avènement des outils d'imagerie diagnostique ;
- un meilleur accès à l'imagerie ;
- une évolution du savoir-faire des cliniciens et radiologues français.

En revanche l'augmentation du nombre d'appendicites compliquées serait imputable à une prise en charge retardée dont l'origine peut être clinique et/ou organisationnelle :

- le patient se présente à un stade précoce avec un tableau clinique incomplet et la prise en charge est reportée. Il est à noter qu'une imagerie réalisée à ce stade peut s'avérer négative car trop tôt pour que la morphologie appendiculaire soit affectée ;
- le patient arrive avec un traitement antalgique ou anti-inflammatoire susceptible de masquer la symptomatologie, le patient est alors renvoyé et il revient quand la situation a évolué vers une complication ;
- dans certaines régions comme il est le cas à Clermont Ferrand, où les plateaux d'Urgence et de chirurgie digestive se situent sur deux sites différents, le transfert des patients d'un site à l'autre rallonge les délais de prise en charge post-diagnostic.

Interrogés sur la variabilité interrégionale évoquée par les demandeurs, les professionnels considèrent que plusieurs facteurs sont susceptibles d'expliquer les écarts et soulignent l'intérêt qu'il y aurait à les documenter. Notamment pour les territoires se situant au niveau des bornes extrêmes par rapport à la moyenne nationale.

Deux facteurs sont d'emblée identifiés par les experts : l'inégalité de l'offre des soins entre les différentes régions et l'existence de facteurs de risques liés au terrain.

En conséquence, l'analyse des écarts de pratiques impose de considérer l'offre de soins, la zone d'attractivité des établissements et l'identification d'éventuels facteurs de risque comme la prévalence de certaines infections, avant de suggérer une imputabilité à des pratiques inappropriées.

6.3 Diagnostic clinique et biologique

Lors de la première réunion du groupe d'orientation, sur la base des recommandations de l'ANDEM de 1996 et de l'expérience pratique des membres du groupe, une discussion préliminaire avait été engagée sur le degré d'actualité des trois éléments ci-dessous, définis par les recommandations de l'ANDEM comme permettant d'exclure le diagnostic d'appendicite grave et de surseoir à l'appendicectomie :

- l'absence de fièvre supérieure à 38° ;
- l'absence de défense localisée à la fosse iliaque droite ;
- l'absence de leucocytose supérieure à 10000.

Le groupe d'orientation a considéré que cette recommandation de l'Andem qui datait de 1996 devait être complétée par l'absence d'augmentation de la CRP.

Le GO a ainsi proposé dans un premier temps la formulation suivante : « En l'absence concomitante de fièvre supérieure à 38°, de défense localisée à la fosse iliaque droite, de leucocytose supérieure à 10000 et d'augmentation de CRP, il n'est pas pertinent de réaliser une appendicectomie en urgence »

Les experts ont considéré qu'à l'opposé, la conjonction des 4 éléments cités ci-dessus pouvait, selon le terrain, représenter un degré de certitude plus ou moins important sur le diagnostic d'appendicite. Chez l'homme jeune, cette tétrade représente une situation de quasi certitude sur le diagnostic. En revanche, chez l'enfant de moins de 15 ans, chez la femme en âge de procréer, chez l'obèse et chez le sujet âgé, la diversité des diagnostics différentiels impose le recours à des examens complémentaires notamment morphologiques

Lors de la seconde réunion du GO, à l'issue de l'analyse critique de la littérature et à la lumière des données recueillies, il a été confirmé qu'aucun signe clinique ou biologique considéré individuellement n'avait de valeur suffisamment discriminante pour permettre de poser le diagnostic d'appendicite.

Ainsi, devant une douleur abdominale, seule la combinaison des trois signes cliniques et biologiques ci-dessous permettait de renseigner sur la probabilité diagnostique d'une appendicite,

- Défense en fosse iliaque droite ;
- Leucocytose supérieure à 10×10^9 cellules/L ;
- CRP supérieure à 8 mg/L.

Contrairement à la première formulation qui reposait sur les seules données de l'ANDEM, la fièvre est apparue, au regard des données recueillies notamment sa valeur diagnostique dans un tableau clinique de suspicion d'appendicite, comme insuffisamment discriminante pour compter parmi les signes cliniques décisionnels.

Concernant les scores clinico-biologiques, les experts soulignent que leur utilisation en France est variable en fonction des équipes.

Le score le plus connu et le plus décrit dans la littérature est le score d'Alvarado. Les données de la littérature ne permettent pas de valider de façon ferme ce score dont le principal biais est sa construction à partir de données rétrospectives recueillies sur des patients déjà opérés pour appendicite.

Certains experts du GO ont cité un score français « le score de François » qui selon les experts semble présenter des résultats intéressants décrits dans une étude prospective publiée en 1991. Cette étude a été analysée par la HAS à la lumière des critères de sélection de la littérature adoptés pour ce travail, elle n'a pas été retenue en raison de sa faible qualité méthodologique.

En conclusion, les experts ne recommandent pas l'utilisation des scores dans la pratique clinique mais considèrent néanmoins que ceux-ci pourraient apporter une aide pour apprécier la probabilité d'appendicite et pour orienter la suite de la prise en charge du patient.

6.4 Les examens d'imagerie

6.4.1 L'ASP

Les membres du groupe d'orientation considèrent qu'il est important de rappeler que conformément aux recommandations de la HAS (26), la radiographie de l'abdomen sans préparation (ASP) n'est pas indiquée pour le diagnostic de l'appendicite

6.4.2 L'échographie

Les experts du groupe d'orientation confirment que les avantages de l'échographie (comparée au scanner) sont : la facilité de réalisation, le moindre coût et l'absence d'irradiation. Les radiologues du groupe d'orientation considèrent néanmoins qu'avant la réalisation de cet examen, il est nécessaire de disposer des éléments cliniques et biologiques permettant d'orienter le processus exploratoire et d'améliorer la performance diagnostique.

Les experts estiment par ailleurs que l'échographie a une bonne valeur diagnostique positive quand l'appendice est visualisé. Ils confirment cependant le caractère opérateur-dépendant « connu » de cet examen susceptible d'affecter sa performance diagnostique.

6.4.3 Le scanner

Selon les experts du GO, l'hétérogénéité des techniques de scanner de même que les modalités de contraste n'influent pas sur la performance diagnostique du scanner. Ce constat est confirmé dans la littérature qui montre que quelles que soient les techniques et les modalités, les données restent homogènes.

Les experts rapportent que dans la pratique française, les modalités de réalisation du scanner sont variables d'un service à l'autre et parfois même au sein d'un même service. Cette variabilité tient du fait que la décision d'injection d'un produit de contraste par le radiologue dépend du contexte

clinique individuel et en particulier de la présence ou pas de graisse intra-abdominale. Cet élément conditionne la qualité de l'examen et son interprétation. Le radiologue est plus à même d'utiliser un produit de contraste chez un patient mince que chez un patient obèse chez qui l'inflammation de la graisse est bien visualisée.

Sur ce point, les données de la littérature ne permettent pas, de trancher sur l'utilisation ou non d'un produit de contraste.

Les experts tiennent tout particulièrement à rappeler que l'imagerie scanner est un examen irradiant et alertent sur le risque du cumul des doses.

Les doses retrouvées dans la littérature et notamment dans la recommandation de l'ACR (30 mSv) sont discutés par les membres du groupe qui considèrent que ces valeurs d'irradiation ne correspondent pas au contexte français, les doses moyennes habituellement utilisées en France sont de l'ordre de 12 mSv (sources IRSN et SFR) et selon l'un des radiologues du groupe les doses seraient plutôt actuellement autour de 3 mSv.

Les experts ont évoqué par ailleurs une publication récente d'une étude randomisée monocentrique coréenne qui compare un scanner à "faibles doses" (2 mSv environ, 116 mGy.cm de valeur médiane) et un scanner standard (8 mSv environ, 521 mGy.cm de valeur médiane). Cette étude conclut, selon les experts, que le scanner à faible dose est non inférieur à un scanner avec une dose standard chez les jeunes adultes avec une suspicion d'appendicite (91).

Les experts soulignent néanmoins que la décision d'utiliser une dose « normale » ou une dose « faible » est prise en fonction de la morphologie du patient, dans le cas d'un patient mince n'ayant pas beaucoup de graisse, il est préférable d'avoir d'emblée une dose « normale » plutôt qu'une dose faible (afin d'éviter le cumul des doses si le premier scanner n'est pas contributif). L'utilisation des scanners à faible dose ne peut par conséquent pas être généralisée.

Concernant le délai de réalisation du scanner en cas de suspicion d'appendicite, les experts se sont référés à un document de consensus, élaboré en 2004 par la Collégiale des radiologues, la Collégiale des urgentistes et la Direction de la politique médicale de l'AP-HP (92). Ce document préconise la réalisation d'un scanner dans les 4h en cas de signes cliniques de gravité (masse à la palpation, empatement...) ou de terrain fragile (corticothérapie, diabète...) et dans un délai de 24h en absence de signes cliniques de gravité.

6.4.4 L'IRM

Cette méthode d'imagerie est utilisée aujourd'hui pour un terrain particulier, la femme enceinte et l'enfant. Dans le champ retenu pour cette évaluation qui exclue ces deux populations, il n'a pas été identifié de données probantes dans la littérature. De plus, compte tenu du parc IRM français et notamment du déficit en matière d'équipements, selon les experts, il ne paraît pas opportun à ce jour et avant amélioration de la situation, d'intégrer cet examen dans la démarche diagnostique d'appendicite pour les autres populations.

6.5 La coelioscopie exploratrice

Au regard des éléments de la littérature analysée, les membres du groupe de travail considèrent que la coelioscopie exploratrice ne doit pas être réalisée en première intention dans un contexte de suspicion d'appendicite, elle peut être proposée chez la femme jeune lorsque l'imagerie (échographie et scanner) n'est pas contributive et qu'il persiste un doute après exclusion de toute de pathologie gynécologique.

Les experts rappellent que la coelioscopie diagnostique est un examen invasif qui présente un taux de morbidité de l'ordre de 2%.

6.6 Eléments décisionnels formalisés sous forme de propositions

Les propositions de référentiels et l'algorithme d'aide à la décision ont été élaborés sur la base des données de la littérature et d'une discussion relative à l'expérience pratique des professionnels du groupe d'orientation qui se sont réunis le 03 mai 2012, le compte rendu in extenso de cette réunion figure en annexe 8.

Les experts du groupe de cotation ont été invités à coter les affirmations sur une échelle de 1 à 9. L'analyse des réponses a été réalisée conformément aux règles de cotation présentées en § 4.2.3. Le tableau d'analyse des réponses figure en **annexe 10**.

Seize experts ont répondu à l'intégralité des questions, un expert (biologiste) avait a priori souhaité répondre exclusivement aux questions concernant son champ de compétences, sa voix n'a été comptabilisée que pour les propositions d'affirmations concernant la biologie.

Neuf propositions sur dix ont été jugées appropriées par les membres du groupe de cotation dont quatre avec un accord fort et cinq avec un accord relatif, une seule proposition a été jugée incertaine.

Proposition 1

Devant une douleur abdominale, aucun signe ou symptôme clinique (y compris la fièvre ou l'apyrexie) n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.

Proposition appropriée avec un accord fort (Cote médiane de 9, minimum 5 et maximum 9)

Proposition 2

Devant une douleur abdominale, aucun paramètre biologique (leucocytes, PNN, CRP) n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.

Proposition appropriée avec un accord fort (Cote médiane de 9, minimum 7 et maximum 9)

Proposition 3

Devant une douleur abdominale évocatrice d'une appendicite, l'absence concomitante d'une défense en fosse iliaque droite, d'une leucocytose supérieure à 10×10^9 cellules/L et d'une CRP supérieure à 8 mg/L rend le diagnostic d'appendicite aiguë peu probable.

Proposition appropriée avec un accord relatif (Cote médiane de 9, minimum 3 et maximum 9)

Proposition 4

Devant une douleur abdominale évocatrice d'une appendicite, la combinaison d'une défense en fosse iliaque droite à une leucocytose supérieure à 10×10^9 cellules/L et une CRP supérieure à 8 mg/L chez l'homme jeune rend le diagnostic d'appendicite aiguë très probable.

Proposition appropriée avec un accord fort (Cote médiane de 9, minimum 7 et maximum 9)

Proposition 5

Au regard de la faible qualité des données publiées, les scores clinico-biologiques d'Alvarado et d'Andersson ne peuvent pas être considérés validés, mais peuvent apporter une aide pour apprécier la probabilité d'appendicite et pour orienter la suite de la prise en charge du patient.

Proposition appropriée avec un accord relatif (Cote médiane de 8, minimum 2 et maximum 9)

Proposition 6

D'après une littérature hétérogène et de faible qualité, l'intérêt diagnostique de l'échographie paraît variable, opérateur et patient dépendant. Lorsque l'échographie visualise l'appendice, sa valeur diagnostique semble être importante. En revanche, la non visualisation de l'appendice ne permet pas d'exclure une appendicite.

Proposition appropriée avec un accord fort (Cote médiane de 8,5, minimum 7 et maximum 9)

Des commentaires ont été émis par deux experts du groupe de cotation : un expert radiologue considère que la formulation de la première phrase « D'après.....et patient dépendant » revêt une connotation très péjorative, qui pourrait dissuader les prescripteurs de commencer par une échographie, néanmoins souvent bien utile en pratique et permettant comme cela est souligné d'éviter le scanner quand elle est positive.

Un autre expert (chirurgien digestif) estime qu'il aurait été intéressant d'inclure également dans la bibliographie des études analysant l'intérêt de l'échographie endovaginale couplée ou non à l'échographie transpariétale en cas de douleur abdominale chez la femme jeune. Il précise, néanmoins que l'échographie endovaginale n'a pas d'intérêt pour le diagnostic positif d'appendicite aiguë mais a une valeur prédictive négative importante en raison de sa performance dans les diagnostics différentiels de la douleur abdominale chez la femme jeune.

Proposition 7

Malgré les biais présents dans les études, les résultats sont concordants et montrent que le scanner est performant pour établir le diagnostic d'appendicite ou pour l'exclure. L'indication de cet examen doit être posée en considérant son caractère irradiant.

Proposition appropriée avec accord relatif (Cote médiane de 9, minimum 6 et maximum 9)

Proposition 8

Le scanner et l'échographie ne sont pas des examens concurrents, ils peuvent avoir chacun leur place dans la démarche diagnostique.

Affirmation jugée incertaine (Cote médiane de 8 : minimum 2 et maximum 9, 2 valeurs inférieures à 5 et 1 manquante)

Les commentaires de deux experts du groupe de cotation (médecin généraliste et gynécologue obstétricien) ont révélé que la formulation de cette proposition a pu prêter à confusion et suggérer le recours systématique aux deux examens d'imagerie.

Il convient par conséquent de préciser, qu'il ne s'agit pas de réaliser les deux examens à la fois mais que le choix de l'une ou l'autre des techniques dépend du patient et de l'environnement. Le recours aux deux techniques est admis uniquement en cas d'une première échographie non concluante.

A ce propos, un autre expert du groupe de cotation (chirurgien digestif) a tenu à préciser que dans le choix de l'un ou l'autre des deux examens, il faudrait prendre en compte, le sexe et le morphotype du patient, la disponibilité des appareils, les antécédents d'examen irradiants, la formation de l'opérateur.

Proposition 9

Il n'y a pas de données de qualité suffisante dans la littérature pour positionner la coelioscopie exploratrice dans la démarche diagnostique.

La coelioscopie exploratrice n'est pas à réaliser en première intention, mais elle peut être proposée lorsque l'imagerie (échographie puis scanner) n'est pas contributive chez la femme jeune.

Proposition appropriée avec accord relatif (Cote médiane de 7, minimum 3 et maximum 9)

Un expert (Gynécologue obstétricien) du groupe de cotation a marqué son désaccord avec la première phrase arguant du fait que les recommandations sont toutes concordantes pour dire que la coelioscopie intervenait en bout de course. Il considère que : « Tout dépend de ce à quoi l'on a accès, s'il n'y a pas accès à l'échographie ou au scanner la coelioscopie est utile, elle se conçoit comme premier geste thérapeutique et devrait être quasi toujours réalisée chez la femme jeune en âge de procréer à qui est suspectée une appendicite »

Un autre expert (Urgentiste) rajoute qu'avant d'avoir recours à la coelioscopie, il faudrait surveiller le patient et répéter l'imagerie.

Un troisième expert (chirurgien digestif) considère quant à lui que la place de la coelioscopie diagnostique ne peut être recommandée et la proposition doit être assortie des données de morbidité et mortalité. Seul le risque d'occlusion sur bride est rapporté. Les risques d'accidents lors de la création du pneumopéritoine doivent également être rapportés et pris en compte dans ce rapport. La place des trocarts unique multiport doit être abordée. Une grande prudence doit être gardée vis à vis de la coelioscopie diagnostique compte tenu des performances des examens complémentaires.

Proposition 10

Au regard des données publiées, le traitement antibiotique ne peut actuellement être considéré validé.

Proposition appropriée avec accord relatif (Cote médiane de 9, minimum 5 et maximum 9)

Un expert du groupe de cotation (chirurgien digestif) avance que ce constat est « confirmé par tous les chirurgiens ».

Un autre expert (gynécologue obstétricien) considère que l'antibiothérapie semble être une bonne alternative étant souvent aussi efficace que l'appendicectomie mais présentant néanmoins plus de récurrence. Il rajoute que pour des personnes à risque chirurgical, obèse ou autre l'antibiothérapie peut être une alternative, sans aggraver les choses

Un troisième expert (médecin généraliste) rappelle l'intérêt potentiel qui ne peut être exclu et considère qu'il faudrait encourager la recherche.

6.7 Proposition d'un arbre décisionnel devant une suspicion d'appendicite

6.7.1 Proposition du groupe d'orientation

Devant une douleur abdominale avec suspicion d'appendicite chez l'adulte, les experts du groupe d'orientation ont défini 3 situations cliniques avec des niveaux de probabilités différents : fort, faible et intermédiaire. Dans ces situations, il est en particulier tenu compte de l'âge, du sexe et du morphotype des patients.

L'arbre décisionnel élaboré par le GO (cf. annexe 7), s'adresse à l'adulte, il ne concerne pas les situations particulières que sont l'appendicite de l'enfant et de la femme enceinte. Il tient compte en particulier des situations de forte prévalence notamment les sujets de moins de 30 ans et les plus de 65 ans.

► Les situations de faible probabilité diagnostique

Devant une douleur abdominale, les situations considérées comme présentant une faible probabilité d'appendicite sont celle où il y a une absence concomitante de la triade suivante :

- Absence de défense abdominale ;
- Absence de leucocytose supérieure à 10.10^9 cellules/L ;
- Absence de CRP supérieure à 8mg/L.

Dans cette situation la conduite à tenir proposée comprend, selon les experts, une surveillance avec une modalité qui devra être adaptée au contexte médical, familial et social, par exemple pour les personnes nomades ou sans domicile fixe, les personnes ayant une phobie du milieu médical, etc..., il est plus prudent de les garder en observation pour éviter une aggravation éventuelle du tableau.

► Les situations de forte probabilité diagnostique

Une seule situation apparaît comme présentant une forte probabilité d'appendicite en présence d'une douleur abdominale, il s'agit de l'association concomitante chez l'homme jeune (âge \leq 30 ans) des trois éléments suivants :

- une défense en FID ;
- une leucocytose $>10.10^9$ cellules/L ;
- une concentration de CRP supérieure à 8mg/L.

Une deuxième situation, également décrite chez l'homme jeune et reposant uniquement sur des signes cliniques est décrite dans la littérature. Il s'agit de la présence concomitante de 3 signes cliniques suivants : une défense en FID, une douleur migrante et une contracture abdominale. Ce tableau très rare par ailleurs (moins de 3% des cas) est selon les experts, plus évocateur d'une péritonite.

► Les situations de probabilité diagnostique intermédiaire ou « doute »

En dehors des situations où les 3 éléments de la triade sont absents (faible probabilité) et la situation où ils sont présents de façon concomitante chez l'homme jeune (forte probabilité), toutes les autres situations relèvent du doute diagnostique et présentent une probabilité intermédiaire. Ces situations nécessitent, selon les experts du groupe d'orientation, des examens complémentaires d'exploration morphologique (échographie ou scanner). Le recours à l'une ou l'autre des deux techniques d'imagerie dépend du terrain.

Chez la femme jeune, il convient de privilégier l'échographie en première intention afin d'éviter l'irradiation et pour exclure une pathologie gynécologique. Si l'échographie n'est pas contributive, il faut faire un scanner. En cas de doute persistant, il faudrait s'orienter vers un avis gynécologique et/ou discuter une laparoscopie exploratrice.

Chez le sujet âgé, le scanner est à envisager car il permet d'éliminer un diagnostic différentiel colique. En cas de scanner négatif il faudrait garder le patient en observation pendant 24h ou l'orienter vers une prise en charge ad-hoc si un autre diagnostic est avéré.

Chez le sujet obèse, la décision de réaliser le scanner ou l'échographie en première intention reste à la discrétion du radiologue. Il a été, toutefois, rappelé les limites de la performance diagnostique de l'échographie dans cette situation. Si l'échographie réalisée en première intention n'est pas contributive, il faudrait réaliser un scanner. Si ce dernier est négatif, il faudrait garder le patient en observation ou l'orienter vers la prise en charge ad-hoc si une autre pathologie est détectée.

Partant de ces différents éléments, les membres du groupe d'orientation ont élaboré en séance un arbre décisionnel qui tient compte des différentes probabilités diagnostiques (cf. annexe 7)

Les experts ont tenu à souligner que ce chemin est envisageable à condition que l'établissement de santé dispose d'un environnement adéquat : plateau technique et professionnels de santé. Or, de l'avis des membres du groupe d'orientation, cela ne semble pas être toujours le cas. C'est la raison pour laquelle en pratique les conditions locales et la problématique organisationnelle peuvent contraindre à adopter une orientation plutôt qu'une autre.

6.7.2 Avis du groupe de cotation sur la proposition d'arbre décisionnel

En deuxième étape de ce travail, il a été recueilli le degré d'accord des experts du groupe de cotation sur l'arbre décisionnel présenté en annexe 7.

A la question : « L'arbre décisionnel tel qu'élaboré vous semble-t-il en accord avec les données de la littérature et votre expérience pratique ? », les experts étaient invités à répondre par **oui, partiellement ou non**, la valeur de cotation attribuée aux différentes réponses étant respectivement de 1, 2 et 3.

A l'issue de l'analyse des réponses, 41% des experts ont répondu oui, 47% ont répondu partiellement et 12% ont répondu non. Globalement 59% des experts sont insatisfaits par l'arbre décisionnel tel que présenté. Les principaux commentaires avancés pour expliquer cet accord « relatif » sont regroupées selon les différentes situations cliniques et selon le degré de probabilité diagnostique :

► Forte probabilité diagnostique

Dans le cas d'une forte probabilité diagnostique, la principale interrogation des experts concerne le recours systématique ou pas à l'imagerie.

Selon un radiologue du groupe de cotation, malgré la forte probabilité positive d'appendicite en présence de la triade défense, hyperleucocytose et CRP chez l'homme jeune, la confirmation du diagnostic par imagerie semble être l'attitude la plus fréquente (actuellement) dans la pratique courante.

Selon un deuxième expert l'algorithme tel qu'élaboré semble être tout à fait en accord avec les données de la littérature et correspond à ce qui devrait être recommandé. Cependant, en pratique, même chez un homme jeune avec la triade présente les chirurgiens ont tendance à demander une imagerie de confirmation (si elle est facilement accessible) avant d'opérer.

Un troisième expert, rajoute que quelque soit la clinique ou la biologie, une technique d'imagerie est toujours pratiquée à la demande des chirurgiens, avant une éventuelle chirurgie.

Un quatrième expert, s'interroge sur la place de la biologie en tant qu'examen de première intention, compte tenu du temps qui peut s'écouler entre l'admission du patient et l'obtention des résultats biologiques. Il considère que, dans cette situation de forte probabilité, l'absence d'imagerie, risquerait d'être associée à un risque plus élevé d'appendicectomie négative et de découverte per opératoire d'autres diagnostics. Il étaye ses propos par les deux références bibliographiques (93),(94). Par ailleurs, il rajoute qu'un recours moins fréquent à l'imagerie pourrait également, probablement, expliquer des différences entre les départements. Considérant qu'en général, le recours à l'imagerie augmenterait avec la spécialisation des centres.

► Faible probabilité

Selon un expert du groupe de cotation, en cas d'absence de la triade, il ne faut pas se contenter de surveiller. S'il existe une douleur exquise à un autre site, l'imagerie est à discuter, même si les leucocytes et la CRP sont normaux car il faut rappeler que l'appendicite n'est pas toujours en fosse iliaque droite, elle peut être atypique (position sous hépatique ou derrière la vessie).

Un autre expert, considère que dans cette situation (absence de triade), plutôt que la surveillance, il est justifié de réaliser des examens complémentaires notamment le scanner compte tenu des coûts actuel de l'hospitalisation comparés aux coûts et la performance de l'imagerie.

Un troisième expert considère que si le contexte est évocateur, la même attitude devrait être adoptée dans les situations de faible probabilité diagnostique "Absence de la triade" et les situations de probabilité intermédiaire "si seulement un ou deux éléments sont présents". De cette formulation, il faudrait par conséquent déduire, un recours systématique à l'imagerie avec une modalité tenant compte de l'âge, du sexe et du morphotype.

Un quatrième expert considère qu'avec les éléments dont nous disposons à ce stade, la probabilité d'une appendicite est faible. Cependant, il s'interroge sur le diagnostic de substitution du patient, qui est venu spontanément ou qui a été adressé par son médecin pour une symptomatologie douloureuse ? La surveillance, en pratique, signifie souvent le retour vers le médecin compte tenu de la fréquentation des services d'urgence. Or, quelle doit être cette surveillance : clinique, biologique, radiologique ? La question sous jacente est de savoir, dans cette situation, quel pourcentage de patients sortira en ayant eu qu'une « prise de sang » et aucun examen spécifique du flanc droit.

Il rajoute qu'étant donné que les résultats des examens biologiques ne sont pas discutables, toute la prise en charge va reposer donc sur la présence ou non d'une défense. Or, la pratique de l'échographie montre (presque quotidiennement) le décalage qui peut exister entre ce qui est indiqué sur une demande d'examen et ce qui est observé par les radiologues. L'examen clinique est-il opératoire dépendant et dans ce cas qui doit prescrire les examens radiologiques : junior, urgentiste, chirurgien ?

Cet expert considère, par ailleurs, que si l'on se base uniquement sur cette triade, il n'est pas tenu compte des appendicites inhabituelles de topographie anormale.

► Probabilité intermédiaire

Dans les situations de probabilité intermédiaire, un expert du groupe de cotation considère que lorsqu'un ou deux éléments sont présents, il est inutile de multiplier les examens complémentaires, L'échographie devrait être réservée aux femmes en période d'activité génitale ou en cas d'examen gynécologique pathologique chez les plus âgées. Chez les autres patients, l'attitude est : un traitement symptomatique avec reconvoque en consultation si le contexte évoque une autre pathologie, ou un scanner directement en précisant "avec opacification digestive ".

Un autre expert considère quant à lui que selon son expérience et selon la littérature, l'absence d'un des signes de la triade doit obliger à donner un autre diagnostic en s'aidant éventuellement des examens complémentaires pour ne pas méconnaître une appendicite débutante due à des germes spécifiques (tels oxyures, shigelles, ...) et risquer de laisser évoluer une appendicite "batarde" vers une péritonite grave.

► Sujets obèses et échographie

Le recours à l'échographie chez le sujet obèse dans les situations de probabilité diagnostique intermédiaire a suscité les commentaires de deux experts du groupe de cotation qui considèrent qu'il n'y a aucun intérêt à effectuer une échographie chez les patients obèses vu la performance chez cette catégorie de patient. Chez le sujet obèse, le scanner devrait être l'examen de première intention quand un ou deux éléments de la triade sont présents. En somme, chez le sujet obèse, ce qui correspond à une douleur abdominale de diagnostic incertain relève d'un scanner conformément aux recommandations de la HAS de 2009 (26).

Un troisième expert considère que chez le sujet obèse, il convient tout d'abord de définir l'indice de masse corporel. Il estime, que les résultats de l'échographie ne sont pas toujours corrélés à l'index de masse corporelle, il s'appuie pour cela sur l'étude de van Randen et al. (95) qui conclue que « la sensibilité de l'échographie n'est pas été significativement affectée par les caractéristiques du patient ou par l'expérience du lecteur » il considère en conséquence que dans cette situation, contrairement à l'affirmation 8, l'échographie et le scanner sont mis en concurrence.

► Cas de la femme jeune

D'après un expert du groupe de cotation, il faudrait s'assurer dans le cas de la femme jeune des « signes négatifs » gynécologiques ou urologiques :

- Absence de grossesse ou de désir de grossesse ;
- absence de leucorrhées ;
- absence de mobilisation utérine douloureuse et de fièvre.

Si ces signes négatifs ne sont pas présents avec la triade, alors c'est une appendicite jusqu'à preuve du contraire et l'échographie pelvienne systématique est recommandée. Cet expert rajoute, en revanche, que si les signes gynécologiques ou d'autres signes sont présents, ou s'il y a une probabilité intermédiaire, l'échographie permettra au chirurgien de différencier entre une pathologie gynécologique ou digestive. En l'absence d'une orientation échographique, il faudra faire une coelioscopie si la défense est persistante malgré la prise d'antalgiques ou de traitement antibiotique et s'il n'y a pas de défense couvrant les différents diagnostics possibles.

Concernant la démarche décisionnelle globale, un expert a décrit l'attitude adoptée en pratique dans son centre, elle qui consiste à faire de l'examen clinique l'élément central de la prise en charge : si l'examen est compatible avec un diagnostic d'appendicite (pas uniquement une défense), une échographie est réalisée. Si l'échographie n'est pas concluante, le patient est réévalué avec les résultats biologiques (scores clinico-biologiques ?) afin de juger de l'indication d'un scanner.

► Conclusion

En définitive, il ressort de l'analyse des différents commentaires des experts du groupe de cotation :

- une tendance chez les experts au recours à l'imagerie, devant une douleur abdominale et suspicion d'appendicite, quelle que soit le degré de probabilité diagnostique ;
- les avis sont partagés quant à l'utilisation de l'échographie ou le scanner en première intention chez le sujet obèse ;
- le recours systématique à l'échographie en première intention chez la femme jeune quel que soit le degré de probabilité diagnostique ; afin d'éliminer les diagnostics différentiels essentiellement gynécologiques ;
- la nécessité de préciser sur l'arbre décisionnel l'absence ou la présence de signes relatifs à une pathologie gynécologique chez la femme jeune.

Conclusion générale

L'objectif du présent travail est de définir les critères d'aide à la décision pour la réalisation d'une appendicectomie. Ces critères s'appuyant sur la performance diagnostique des éléments cliniques, biologiques et des examens morphologiques (imagerie et coelioscopie diagnostique) et sur l'avis des professionnels.

Plusieurs éléments de conclusion peuvent être retenus de ce travail.

Depuis les recommandations de l'ANDEM de 1996, deux éléments nouveaux ont pu être identifiés comme apportant une valeur ajoutée au processus diagnostique, il s'agit de la CRP et du scanner.

Par ailleurs, sur la base des données de la littérature les professionnels de santé ont souligné l'insuffisance des valeurs diagnostiques individuelles des signes et symptômes cliniques et des paramètres biologiques.

Les recommandations de l'ANDEM de 1996 avaient défini un faisceau d'éléments (une fièvre supérieure à 38°, une défense en fosse iliaque droite et une leucocytose supérieure à 10×10^9 cellules/L) qui permettaient quand ils étaient absents de surseoir à une appendicectomie. Les données actuelles, confortées par les professionnels de santé ont permis d'identifier la CRP comme paramètre biologique apportant une valeur ajoutée au processus diagnostique et d'exclure la fièvre considérée au regard des données de la littérature comme très peu discriminante.

Ainsi les professionnels considèrent qu'il convient aujourd'hui de considérer, devant une douleur abdominale aiguë, les éléments de la triade suivante « **une défense en fosse iliaque droite, une leucocytose supérieure à 10×10^9 cellules/L et une CRP supérieure à 8 mg/L** » pour définir trois niveaux de probabilité diagnostique (faible, intermédiaire et forte).

Ces niveaux de probabilité, reposent sur l'absence ou la présence totale ou partielle des éléments de la triade et peuvent constituer une aide à la décision pour le recours aux examens complémentaires notamment l'imagerie.

Concernant l'imagerie, les experts ont tenu à ce que soit souligné dans le rapport d'évaluation que l'ASP n'était pas à réaliser.

Par ailleurs, dans le choix de la technique d'imagerie, le morphotype, le sexe, l'âge des patients et l'aspect irradiant de la technique, doivent être considérés.

Ainsi, lorsque l'imagerie est nécessaire, l'échographie est l'examen à considérer en première intention, chez le sujet masculin jeune et chez la femme. Elle possède une valeur diagnostique importante en cas de visualisation de l'appendicite ce qui en fait une exploration intéressante, car elle permettrait ainsi d'éviter le scanner dont le caractère irradiant et le risque de cumul d'exposition (en particulier chez le sujet jeune) ont été rappelés. Il a été toutefois souligné qu'il s'agissait d'un examen opérateur et patient dépendant (sexe et morphotype).

Chez le sujet âgé, l'utilisation du scanner est préconisée, l'examen apparaissant plus approprié pour établir un diagnostic différentiel colique.

Chez le sujet obèse, les avis des professionnels sont partagés quant à l'utilisation en première intention du scanner ou de l'échographie.

En règle générale, dans les situations où l'échographie n'est pas contributive, un recours au scanner est préconisé.

Selon les experts, le délai de réalisation d'un scanner est de 4h en cas de signes cliniques de gravité (masse à la palpation, empatement...) ou de terrain fragile (corticothérapie, diabète) et dans un délai de 24h en absence de signes cliniques de gravité.

Concernant l'exploration coelioscopique, les éléments de la littérature et les professionnels de santé permettent de considérer, qu'avant d'être envisagée chez la femme jeune, les autres

moyens d'exploration morphologique que sont l'échographie et le scanner doivent avoir été épuisés et un avis gynécologique doit avoir été demandé.

Ces différents éléments ont été structurés, avec la contribution des membres du groupe d'orientation, sous la forme d'un schéma d'aide à la décision, qui a été soumis au groupe de cotation. La principale proposition qui ne semble pas rallier tous les experts est l'absence de recours systématique à l'imagerie dans les situations de faible et de forte probabilité diagnostique, certains membres du groupe de cotation considérant nécessaire de ne pas exclure la possibilité de la réaliser.

LA HAS considère que ce schéma, peut contribuer à clarifier la démarche diagnostique en catégorisant le degré de probabilité d'une appendicite en fonction des paramètres cliniques et biologiques proposés, le recours à l'imagerie (échographie à minima), restant à la discrétion des praticiens, en particulier dans les situations de faible et forte probabilité diagnostique.

Concernant le traitement antibiotique des appendicites aiguës non compliquées, la HAS considère que son efficacité n'a pas encore été démontrée de façon significative pour permettre une substitution aujourd'hui à l'appendicectomie. Des études contrôlées randomisées sont nécessaires pour démontrer l'efficacité des antibiotiques notamment à long terme et pour identifier les sous-groupes de patients pouvant bénéficier de ce traitement, une étude française est prévue.

Sur la base de l'ensemble des éléments recueillis après une analyse critique des données de la littérature et après consultation des experts, la HAS conclut :

- Devant une douleur abdominale, aucun signe ou symptôme clinique (y compris la fièvre ou l'apyrexie) n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.
- Devant une douleur abdominale, aucun paramètre biologique (leucocytes, PNN, CRP) n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.
- Devant une douleur abdominale évocatrice d'une appendicite, l'absence concomitante d'une défense en fosse iliaque droite, d'une leucocytose supérieure à 10×10^9 cellules/L et d'une CRP supérieure à 8 mg/L rend le diagnostic d'appendicite aiguë peu probable.
- Devant une douleur abdominale évocatrice d'une appendicite, la combinaison d'une défense en fosse iliaque droite à une leucocytose supérieure à 10×10^9 cellules/L et une CRP supérieure à 8 mg/L chez l'homme jeune rend le diagnostic d'appendicite aiguë très probable.
- Les scores clinico-biologiques d'Alvarado et d'Andersson ne peuvent pas être considérés validés, mais peuvent apporter une aide pour apprécier la probabilité d'appendicite et pour orienter la suite de la prise en charge du patient.
- L'intérêt diagnostique de l'échographie paraît variable, opérateur et patient dépendant. Lorsque l'échographie visualise l'appendice, sa valeur diagnostique semble être importante. En revanche, la non visualisation de l'appendice ne permet pas d'exclure une appendicite.
- le scanner est performant pour établir le diagnostic d'appendicite ou pour l'exclure. L'indication de cet examen doit être posée en considérant son caractère irradiant.

- La coelioscopie exploratrice n'est pas à réaliser en première intention, elle peut être proposée chez la femme jeune lorsque l'imagerie (échographie puis scanner) n'est pas contributive et après avis gynécologique.
- Le traitement antibiotique de l'appendicite ne peut actuellement être considéré validé.

La HAS tient à rappeler par ailleurs que :

L'ASP n'est pas indiquée pour le diagnostic de l'appendicite.

HAS/Radiographie de l'abdomen sans préparation-rapport d'évaluation janvier 2009

Annexe 1. Paramètres de l'évaluation diagnostique : quelques définitions

Courbe ROC : L'étude de la courbe ROC (Receiver Operating Characteristic) permet d'argumenter le choix du seuil diagnostique d'un test en recherchant le meilleur rapport entre sensibilité et spécificité de ce test

Sensibilité : la sensibilité d'un test ou d'un examen diagnostique est sa capacité à donner un résultat positif lorsque la maladie (ou la condition) est présente,

$$Se = a / (a+c)$$

Spécificité : la spécificité d'un test ou d'un examen diagnostique est sa capacité à donner un résultat négatif lorsque la maladie (ou la condition) n'est pas présente,

$$Sp = d / (b+d)$$

Les valeurs prédictives : elles expriment comment les résultats d'un examen diagnostique vont prédire la présence ou l'absence d'une maladie.

La valeur prédictive positive (VPP) : elle exprime la probabilité que la maladie recherchée soit effectivement présente chez le sujet ayant le test positif

$$VPP = a / (a+b)$$

La valeur prédictive négative (VPN) : elle exprime la probabilité que la maladie recherchée n'est pas présente chez le sujet ayant le test négatif

$$VPN = d / (c+d)$$

L'efficacité diagnostique ou précision diagnostique : est obtenue en divisant les réponses correctes (vrais positifs et vrais négatifs) par le nombre total de patients

$$(a+d) / (a+b+c+d)$$

	Malades	Non-malades
Test positif	a (vrais positifs)	b (faux positifs)
Test négatif	c (faux négatifs)	d (vrais négatifs)

Les rapports de vraisemblance (" likelihood ratios ")

Les rapports de vraisemblance ou Likelihood ratio, sont un autre mode d'expression des caractéristiques intrinsèques d'un test. Ils estiment le rapport entre la probabilité d'avoir un test positif (ou négatif) chez les sujets malades et celle d'avoir un test positif (ou négatif) chez les sujets sains. Calculés à partir de la Se et de la Sp du test, ils sont donc indépendants de la prévalence de la maladie dans la population.

Le rapport de vraisemblance positif (RVP) est égal au taux de tests positifs chez les malades (soit la Se) sur le taux de tests positifs chez les non malades soit $[1 - Sp]$:

$$RVP = Se / (1 - Sp)$$

Il quantifie le gain diagnostique d'un test positif, un individu malade ayant x fois plus de chance d'avoir un test positif qu'un individu sain.

Le rapport de vraisemblance négatif (RVN) est égal au taux de tests négatifs chez les malades soit $[1 - Se]$ sur le taux de tests négatifs chez les sujets sains soit la Sp :

$$RVN = (1 - Se)/Sp$$

Il quantifie le gain diagnostique d'un test négatif, un individu malade ayant x fois moins de chance d'avoir un test négatif qu'un individu sain.

Plus le rapport de vraisemblance positif est élevé et plus le rapport de vraisemblance négatif est faible, plus le gain diagnostique du test est important

Apport diagnostique d'un test en fonction de la valeur des rapports de vraisemblance positif et négatif (96)

RVP	RVN	Apport diagnostique
>10	< 0,1	Très fort
5-10	0,1-0,2	Fort
2-5	0,2-0,5	Modéré
1-2	0,5-1	Faible
1	1	Nul

Annexe 2. Codes GHM utilisés pour la recherche des données de pratique

Codes utilisés pour le recueil des données de 1999 à 2008

GHM version 10b CMD 06 affection du tube digestif

CMD 06 : affections du tube digestif

N°GHM	Libelle GHM
06C08W	Appendicectomies compliquées, âge supérieur à 69 ans et/ou CMA
06C08V	Appendicectomies compliquées, âge inférieur à 70 ans sans CMA
06C09W	Appendicectomies non compliquées, âge supérieur à 69 ans et/ou CMA
06C09V	Appendicectomies non compliquées, âge inférieur à 70 ans sans CMA

Codes utilisés pour le recueil des données de l'année 2009

Modification des GHM février 2009

GHM - version 11

CMD 06 : affections du tube digestif

N°GHM	Libelle GHM
06C08	Appendicectomies compliquées
06C081	Appendicectomies compliquées, niveau 1
06C082	Appendicectomies compliquées, niveau 2
06C083	Appendicectomies compliquées, niveau 3
06C084	Appendicectomies compliquées, niveau 4

* Dans la version 11, l'appendicite avec complication est recherchée en position de diagnostic principal ou associé, plus en position de diagnostic relié.

* Dans la version 11, cette racine est éligible aux niveaux de sévérité. L'âge intervient pour les niveaux de sévérité : un âge supérieur à 79 ans classe un RSS de niveau 1 en niveau 2, un RSS de niveau 2 en niveau 3 et un RSS de niveau 3 en niveau 4. Elle nécessite une confirmation de codage pour les séjours sans nuitée (0 jour) sans mode de sortie décès ou transfert vers le MCO. Attention, pour cette racine, il existe des exclusions GHM/CMA (liste D-094).

Manuel des GHM - Version n°11 - 2009 06-10

06C08 Appendicectomies compliquées

Voir la liste A-047 : Appendicectomies

Voir la liste D-094 : Appendicites avec complications

N°GHM	Libelle GHM
06C09	Appendicectomies non compliquées
06C091	Appendicectomies non compliquées, niveau 1
06C092	Appendicectomies non compliquées, niveau 2
06C093	Appendicectomies non compliquées, niveau 3
06C094	Appendicectomies non compliquées, niveau 4

Annexe 3. La stratégie de recherche et la liste des sources interrogées

Stratégie d'interrogation des bases bibliographiques et résultats

La stratégie d'interrogation des bases de données précise pour chaque question et / ou types d'étude les termes de recherche utilisés, les opérateurs booléens et la période de recherche.

Les termes de recherche utilisés sont soit des termes issus de thésaurus (descripteurs), soit des termes libres (du titre ou du résumé). Ils sont combinés avec les termes décrivant les types d'études.

La recherche a porté sur les publications en langue anglaise et française.

Le tableau 1 présente de façon synthétique les étapes successives de cette interrogation.

Le nombre total de références obtenues est 759 (au 04/04/12).

Stratégie d'interrogation documentaire dans la base Medline.

Type d'étude / sujet		Période	Nombre de références
Termes utilisés			
Diagnostic de l'appendicite			
Recommandations		Pas de limite – 03/2011	14
Etape 1	(Appendicitis/de AND Diagnosis!/de) OR Appendicitis/diagnosis/de OR appendicitis/ti,ab		
ET			
Etape 2	Health Planning Guidelines/de OR (Practice Guideline OR Guideline OR Consensus Development Conference OR Consensus Development Conference, NIH)/pt OR (recommendation* OR guideline*)/ti		
Méta-analyses, revues systématiques		Pas de limite – 03/2011	63
Etape 1			
ET			
Etape 3	Meta-Analysis/pt OR systematic review/ti,ab OR (metaanalysis OR meta analysis)/ti		
Douleur abdominale et appendicite			
Recommandations		Pas de limite – 10/2011	2
Etape 4	(Abdominal Pain/de AND Appendicitis/de) OR (abdominal pain/ti,ab AND appendi*/ti,ab		
ET			
Etape 2			
Méta-analyses, revues systématiques		Pas de limite – 10/2011	8
Etape 4			
ET			
Etape 3			

Diagnostic de l'appendicite par imagerie médicale		
Essais contrôlés randomisés		Pas de limite – 42 12/2011
Etape 5	(Appendectomy OR Appendicitis)/de OR (appendicectom* OR appendectom* OR appendicitis)/ti,ab	
ET		
Etape 6	Diagnostic Imaging!/de OR (magnetic resonance* OR ultrasound OR ultrasonogra* OR sonogra* OR tomography* OR imaging OR radiogram* OR X-ray)/ti,ab OR (CT OR MRI OR MDCT)/ti	
ET		
Etape 7	(Random Allocation OR Double-Blind Method OR Single-Blind Method OR Cross-Over Studies)/de OR Randomized Controlled T/pt OR random*/ti,ab	
Diagnostic de la douleur abdominale par imagerie médicale		
Essais contrôlés randomisés		Pas de limite – 115 12/2011
Etape 8	Abdominal Pain/de OR abdominal pain/ti,ab	
ET		
Etape 7		
Diagnostic biologique ou examen physique de l'appendicite		
Essais contrôlés randomisés		01/1990 – 88 01/2012
Etape 9	Appendicitis/diagnosis/de OR (appendicitis/ti AND (diagnosis OR examination)/ti)	
ET		
Etape 10	(Random Allocation OR Double-Blind Method OR Single-Blind Method OR Cross-Over Studies)/de OR Randomized Controlled Trial/pt OR random*/ti,ab	
SAUF	Etape 6	
Etudes observationnelles		01/1990 – 287 01/2012
Etape 9		
ET		
Etape 11	(Cohort Studies OR Longitudinal Studies OR Follow-Up Studies OR Prospective Studies OR Clinical Trials as Topic OR Case-Control Studies OR Retrospective Studies)/de OR (Comparative Study OR Clinical Trial)/pt OR (cohort study OR cohort studies OR clinical trial OR retrospective stud* OR comparative stud* OR versus)/ti	
SAUF	Etape 6	

Diagnostic biologique ou examen physique lors de douleurs abdominales		
Essais contrôlés randomisés	01/1990 – 01/2012	45
Etape 12 Abdominal pain/diagnosis/de OR (abdominal pain/ti AND (diagnosis OR examination)/ti)		
ET		
Etape 10		
SAUF Etape 6		
Etudes observationnelles	01/1990 – 01/2012	89
Etape 12		
ET		
Etape 11		
SAUF Etape 6		

* troncature ; de : descriptor ; ti : title ; ab : abstract ; pt : publication type ; ! : Explosion

Liste des sites consultés et résultats

Agence française de sécurité sanitaire des produits de santé

Bibliothèque médicale Lemanissier

Catalogue et index des sites médicaux francophones – CISMef

Comité d'Evaluation et de Diffusion des Innovations Technologiques – CEDIT

Evaluation des technologies de santé pour l'aide à la décision (Fédération hospitalière de France) – ETSAD

Expertise collective INSERM

Fédération de chirurgie viscérale et digestive

Société française d'endoscopie digestive

Société française de médecine générale – SFMG

Société française de pathologie

Société française de radiologie – SFR

Société d'imagerie abdominale et digestive

Société nationale française de gastro-entérologie – SFNGE

Adelaide Health Technology Assessment – AHTA

Agency for Healthcare Research and Quality – AHRQ

Alberta Heritage Foundation for Medical Research – AHFMR

Alberta Medical Association

American College of Gastroenterology

American College of Physicians – ACP

American Gastroenterological Association – AGA

American Society of Colon and Rectal Surgeons

Australian Safety and Efficacy Register of New Interventional Procedures – Surgical
Blue Cross Blue Shield Association – BCBS – Technology Evaluation Center
BMJ Clinical Evidence
California Technology Assessment Forum – CTAF
Canadian Agency for Drugs and Technologies in Health – CADTH
Canadian Task Force on Preventive Health Care
Centers for Disease Control and Prevention – CDC
Centre fédéral d'expertise des soins de santé – KCE
Centre for Clinical Effectiveness – CCE
Centre for Reviews and Dissemination databases
Clinical Knowledge Summaries
Clinical Practice Guidelines Portal
CMA Infobase
Cochrane Library
College of Physicians and Surgeons of Alberta – CPSA
European Association for Endoscopic Surgery
European Society of Gastrointestinal and Abdominal Radiology
European Society of Gastrointestinal Endoscopy
Euroscan
Guideline Advisory Committee – GAC
Guidelines and Protocols Advisory Committee- GPAC
Guidelines International Network - GIN
Health Services Technology Assessment Text – HSTAT
Horizon Scanning
Institut national d'excellence en santé et en services médicaux – INESS
Institute for Clinical Evaluative Sciences – ICES
Institute for Clinical Systems Improvement – ICSI
Institute for Health Economics Alberta – IHE
International Network of Agencies for Health Technology Assessment – INAHTA
Intute Health & Life Sciences – INTUTE
Medical Services Advisory Committee – MSAC
National Coordinating Centre for Health Technology Assessment – NCCHTA
National Guideline Clearinghouse – NGC
National Health and Medical Research Council – NHMRC
National Horizon Scanning Centre – NHSC
National Institute for Health and Clinical Excellence - NICE
New Zealand Guidelines Group – NZGG

New Zealand Health Technology Assessment – NZHTA

NHS Evidence

Ontario Health Technology Advisory Committee – OHTAC

Public Health Agency of Canada - Diseases Prevention and Control Guidelines

Scottish Intercollegiate Guidelines Network – SIGN

Singapore Ministry of Health

Society for Surgery of the Alimentary Tract

Tripdatabase

U.S. Preventive Services Task Force

Veterans Affairs Technology Assessment Program

Veterans affairs, Dep. Of Defense Clinical practice guidelines

West Midlands Health Technology Assessment Collaboration – WMHTA

World Gastroenterology Organisation

Annexe 4. Grille AMSTAR

	Questions	Réponses
1	<p>A-t-on fourni un plan « à priori » ?</p> <p>La question à l'étude et les critères d'inclusion devraient être établis avant l'exécution de l'examen systématique.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
2	<p>Y avait-il un double moyen de choisir le sujet d'analyse et d'extraire les données ?</p> <p>Il devrait y avoir au moins deux extracteurs de données indépendants et un mécanisme pour arriver à un consensus dans les cas de divergences.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
3	<p>A-t-on effectué une recherche complète dans la littérature ?</p> <p>La recherche devrait porter sur au moins deux sources électroniques. Le rapport doit inclure les années et les bases de données utilisées (ex. : Central, EMBASE et MEDLINE). Les auteurs doivent fournir les mots clés et/ou les termes de la chaîne utilisés et, lorsque cela est possible, la stratégie de recherche. Toutes les recherches doivent être complétées par une consultation des contenus courants, des revues, des manuels, de registres spécialisés ou de spécialistes du domaine à l'étude, et par une revue des références contenues dans les études.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
4	<p>Est-ce que la nature d'une publication (ex. : littérature grise) a servi de critère d'inclusion ?</p> <p>Les auteurs devraient déclarer qu'ils ont cherché des rapports d'études sans égard au type de publication. Ils devraient aussi dire s'ils ont exclu des rapports à cause de la nature de la publication, de sa langue, etc.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
5	<p>Les auteurs devraient fournir la liste des études incluses et des études exclues</p> <p>Les auteurs devraient fournir la liste des études incluses et des études exclues.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
6	<p>Les auteurs ont-ils fourni une description des caractéristiques des études incluses ?</p> <p>Présentées sous une forme condensée comme un tableau, les données de l'étude originale devraient inclure les participants, les interventions et les résultats. L'étude devrait rendre compte des différentes caractéristiques de toutes les études analysées (ex. : âge, race, sexe, données socioéconomiques pertinentes, état de la maladie, durée, sévérité ou autres maladies).</p>	<p>Oui Non Ne peut répondre Sans objet</p>

7	<p>La qualité scientifique des études incluses dans l'examen a-t-elle été analysée et documentée ?</p> <p>Les méthodes d'évaluation a priori devraient être fournies (pour les études d'efficacité si l'auteur a choisi de n'inclure que les études aléatoires, les essais à double insu, les essais comparatifs avec placebo, ou l'allocation dissimulée utilisée comme critère d'inclusion); pour d'autres types d'études, des éléments différents pourront être pertinents.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
8	<p>La qualité scientifique des études incluses a-t-elle été utilisée de façon appropriée dans la formulation des conclusions ?</p> <p>Les résultats au chapitre de la rigueur méthodologique et de la qualité scientifique devraient être pris en compte dans l'analyse et les conclusions de l'examen systématique, et devraient être mentionnés explicitement dans la formulation des recommandations.</p>	<p>Oui Non Ne peut répondre Sans objet</p>
9	<p>Les méthodes de groupement des résultats des études étaient-elles appropriées ?</p> <p>Lorsqu'on regroupe des résultats, on devrait d'abord vérifier si les études sont combinables en appliquant un test d'homogénéité (p. ex., le test I2 qui détermine l'homogénéité des données). S'il existe de l'hétérogénéité, on devrait utiliser un modèle d'analyse des effets aléatoires et/ou considérer la pertinence de regrouper les résultats des études (est-il approprié de les regrouper ?).</p>	<p>Oui Non Ne peut répondre Sans objet</p>
10	<p>A-t-on analysé la possibilité d'un biais de publication ?</p> <p>L'analyse d'un biais de publication devrait inclure des représentations graphiques (p. ex., graphique en entonnoir ou autres tests disponibles) et/ou des analyses statistiques (p. ex. : le test de régression de Egger).</p>	<p>Oui Non Ne peut répondre Sans objet</p>
11	<p>A-t-on déclaré les conflits d'intérêts ?</p> <p>Les sources potentielles de soutien devraient être clairement reconnues dans les examens systématiques et dans les études incluses dans ces examens.</p>	<p>Oui Non Ne peut répondre Sans objet</p>

D'après la grille publiée par le CCNMO (Centre de Collaboration Nationale des Méthodes et Outils) dernière mise à jour 27 mars 2012

<http://www.nccmt.ca/index-fra.html>

Annexe 5. Compte rendu de la première réunion du groupe d'orientation du 15 décembre 2011

APPENDICECTOMIE

Éléments décisionnels pour une indication pertinente

Membres du groupe d'orientation

Membres présents : Dr Hervé BERCHE, Pr Philippe MARTEAU, Pr Christophe PONCELET, Pr Karem SLIM, Dr Nathalie SIAUVE, Pr Bertrand TOUSSAINT

Membres absents : Dr Isabelle KAUFFMANN

Représentants de la HAS /SEAP :

Dr Sun Hae Lee Robin, Dr Michèle Morin Surroca, Dr Nadia Zeghari-Squalli PhD.

Objet : Identifier et discuter des questions à traiter pour ce sujet, au regard de l'état de l'art et des difficultés rencontrées dans la pratique professionnelle

En préambule certains membres du groupe d'orientation se sont interrogés sur la définition de la pertinence en fonction de quel point de vue. Il a été rappelé que la qualité des soins étant au centre des missions de la HAS, le concept de pertinence pouvait être défini comme suit : juste soin au bon patient, au bon moment au coût acceptable. Les définitions de la RAND et du NHS ont été rappelées.

Discussion des données de pratiques présentées dans le document de cadrage

A propos de l'augmentation du taux d'appendicectomies compliquées sur la période 1999-2009

Les membres du groupe d'orientation notent qu'il serait intéressant d'exprimer les chiffres concernant l'augmentation des appendicectomies compliquées (en fonction des années) en valeur absolue plutôt qu'en pourcentage par rapport à la valeur référence de 1999. Ils estiment qu'il serait également nécessaire que ces chiffres soient croisés avec les durées de prise en charge préopératoire. La faisabilité de cette requête sur les données PMSI et CCAM sera étudiée.

A propos de la variabilité interrégionale

Les experts ont proposé plusieurs facteurs susceptibles d'expliquer les écarts et soulignent l'intérêt qu'il y aurait à les documenter. Notamment pour les territoires se situant au niveau des bornes extrêmes par rapport à la moyenne nationale.

Par exemple, pour la région du Nord-Pas de Calais identifiée comme « à forte incidence » la récupération par cette région des patients des territoires avoisinants, frappés par une forte désertification médicale pourrait constituer un des motifs. A l'opposé, la Lorraine enregistre le taux national le plus bas, mais il s'agit d'une région touchée par une forte désertification médicale.

L'exemple de la Corse a été également cité. Il s'agit d'une région présentant le taux le plus élevé d'appendicectomies sur le plan national, toutefois, c'est une zone d'endémie en oxyuroses, facteur favorisant l'obstruction de la lumière appendiculaire et par conséquent la survenue d'appendicite aiguë.

Deux facteurs sont d'emblée identifiés par les experts : l'inégalité de l'offre des soins entre les différentes régions et l'existence de facteurs de risques liés au terrain.

En conséquence, l'analyse des écarts de pratiques impose de considérer l'offre de soins, la zone d'attractivité des établissements et l'identification d'éventuels facteurs de risque comme la prévalence de certaines infections, avant de suggérer une imputabilité à des pratiques inappropriées.

Un facteur tarifaire a également été évoqué. L'exploration diagnostique par coelioscopie étant moins valorisée que la réalisation d'une appendicectomie par coelioscopie, certains professionnels pourraient être amenés à réaliser dans certains établissements, une appendicectomie systématique dans la foulée. Il serait intéressant à cet égard de comparer l'évolution des appendicectomies par laparotomie à celles réalisées par coelioscopie.

Pertinence d'une appendicectomie : rappel de l'objectif du travail

L'objectif de ce travail est d'identifier les patients pour lesquelles il est pertinent de réaliser une appendicectomie. Le groupe d'orientation insiste sur la définition de la sémantique qui semble importante dans ce contexte, il s'agit de répondre précisément à ces deux questions

- *Est-ce que c'est une appendicite aiguë ?*
- *Faut-il réaliser une appendicectomie ou pas ?*

Les membres du GO ont souligné l'importance de la pertinence de l'acte chirurgical en rappelant qu'une appendicectomie inutile n'est pas sans risques. L'apparition fréquente de brides intestinales constitue la première cause d'occlusion post-opératoire (2 à 3%) et représente 40% des causes d'occlusions intestinales aiguës. Chez la femme en âge de procréer, l'appendicectomie est l'un des facteurs de risques de la grossesse extra-utérine.

Dans l'évaluation de la pertinence d'une appendicectomie, la prise en compte de l'histoire naturelle de l'appendicite est primordiale selon les experts.

Les formes sont très variables, les appendicites peuvent être graves d'emblée, moins graves ou résolutive.

Le rôle de l'anatomo-cytopathologiste est prépondérant pour définir la signification histopathologique de chaque forme et confirmer ou infirmer une appendicite.

Ceci est d'autant plus important dans le cas des appendicites médicales « formes virales » qui peuvent justifier une appendicectomie blanche.

Il convient de parler des appendicites plutôt que de l'appendicite.

Les questions suivantes méritent donc d'être considérées :

- Analyse de la concordance entre l'aspect histologique de la pièce d'exérèse et le tableau clinique faisant suspecter une appendicite ;
- Caractérisation histopathologique de la pièce d'exérèse.

Compte tenu de ces éléments, la présence d'un anatomopathologiste, au sein du groupe d'orientation, paraît nécessaire pour définir avec précision les critères de classification cytopathologique des appendicites. Il a été décidé à ce titre d'associer un anatomopathologiste à la prochaine réunion du GO.

Pour la réalisation de ce travail, notamment pour l'élaboration d'un processus décisionnel, le GO considère que la première étape clé est de distinguer deux situations cliniques :

- les situations de certitude diagnostique ;
- les situations de doute diagnostique.

Pour cela il faut s'appuyer sur les valeurs prédictives de tous les outils diagnostiques et sur les variables cliniques et biologiques.

Examens clinico-biologiques, place des scores

Les examens clinico-biologiques permettent de stratifier le risque d'appendicite et d'orienter la décision de prise en charge (sortie, surveillance, examens complémentaires ou appendicectomie). Il est nécessaire pour cela de définir la valeur prédictive des variables biologiques et des signes et symptômes cliniques. Les membres du GO rappellent l'intérêt potentiel des scores clinico-biologiques à cette étape du processus diagnostique et considèrent qu'il est important de traiter ce point, même s'ils admettent ne pas l'utiliser en pratique courante.

Aspect morphologique, place de l'imagerie et de la coelioscopie

L'évaluation de la valeur prédictive des examens d'imagerie notamment l'échographie et le scanner, en tenant compte du terrain, est un des points clés du travail.

Le groupe a toutefois souligné que certaines interventions auront lieu sans imagerie, par indisponibilité de plateaux d'imagerie ou d'absence de radiologues, malgré la prise en charge du patient dans un établissement autorisé à accueillir les urgences.

Les membres du groupe insistent également sur l'intérêt de définir la place de l'exploration par coelioscopie dans le diagnostic de l'appendicite. Ils recommandent de consulter les recommandations éditées en la matière par l'European Association for Endoscopic Surgery (EAES). La coelioscopie fait partie du processus diagnostique, notamment chez la femme en âge de procréer.

A propos de la recherche bibliographique et de la littérature identifiée

Une première recherche bibliographique avec une stratégie limitée aux revues systématiques et méta-analyses sans restriction sur la période a été réalisée, il en ressort :

- un très faible corpus de la littérature ;
- des articles avec beaucoup de faiblesses méthodologiques ;
- des articles ne répondant pas de façon directe aux questions posées.

Le groupe d'orientation (GO) a discuté cette stratégie, il estime que compte tenu de la pauvreté de la littérature identifiée avec ces critères, il est important d'inclure les études originales autant que de besoin qui peuvent apporter des éléments supplémentaires non documentés dans les méta-analyses et revues systématiques disponibles.

Questions ayant servi de base à la discussion

Est-ce que la recommandation qui définit les trois éléments ci-dessous comme permettant d'exclure le diagnostic d'appendicite grave et de surseoir à l'appendicectomie est toujours d'actualité ?

- l'absence de fièvre supérieure à 38° ;
- l'absence de défense localisée à la fosse iliaque droite ;
- l'absence de leucocytose supérieure à 10000.

Le groupe d'orientation considère que cette recommandation de l'Andem qui date de 1996 doit être complétée par l'absence d'augmentation de la CRP dont la valeur prédictive négative associée à une absence d'hyperthermie est élevée (près de 100%). Par ailleurs, il considère que les patients pour qui ces 4 éléments ne sont pas concomitants n'ont aucun risque de péritonite dans les 12 -24 h qui suivent et le fait d'attendre n'aggrave pas le pronostic.

Le GO propose ainsi que la formulation de cette recommandation soit modulée de la manière suivante : « En l'absence concomitante de fièvre supérieure à 38°, de défense localisée à la fosse iliaque droite, de leucocytose supérieure à 10000 et d'augmentation de CRP, il n'est pas pertinent de réaliser une appendicectomie en urgence »

Quels sont les éléments cliniques et para-cliniques devant lesquels une indication d'appendicectomie peut être posée ?

A l'opposé, la conjonction des 4 éléments ci-dessous peuvent, selon le terrain, représenter un degré de certitude plus ou moins important sur le diagnostic d'appendicite :

- fièvre supérieure à 38 ;
- défense localisée à la fosse iliaque droite ;
- leucocytose supérieure à 10000 ;
- CRP élevée.

Chez l'homme jeune, cette tétrade représente une situation de quasi certitude sur le diagnostic. En revanche, chez l'enfant de moins de 15 ans, chez la femme en âge de procréer, chez l'obèse et chez le sujet âgé, la diversité des diagnostics différentiels impose le recours à des examens complémentaires notamment biologiques et morphologiques.

Place de l'imagerie dans stratégie diagnostique

- Fiabilité de l'échographie ;
- Fiabilité du scanner (avec ou sans produit de contraste) ;
- Fiabilité de l'IRM.

Selon les membres de groupe d'orientation, l'échographie n'a de valeur que positive, elle est opératoire dépendant mais présente indéniablement les avantages suivants : facilité de réalisation, moindre coût et absence d'irradiation.

La performance diagnostique du scanner est élevée (une sensibilité et une spécificité au-dessus de 90%), son principal avantage est qu'il permet d'explorer toute la cavité abdominale et d'éliminer les diagnostics différentiels mais son principal inconvénient est l'irradiation qui limite son utilisation notamment chez l'enfant, la femme en âge de procréer et la femme enceinte.

Quant à la performance du scanner avec ou sans injection de produit de contraste, le groupe d'orientation estime qu'il n'y a pas de recommandation à émettre sur ce point, le choix d'utiliser ou pas le produit de contraste dépend essentiellement du terrain et de l'expérience des professionnels.

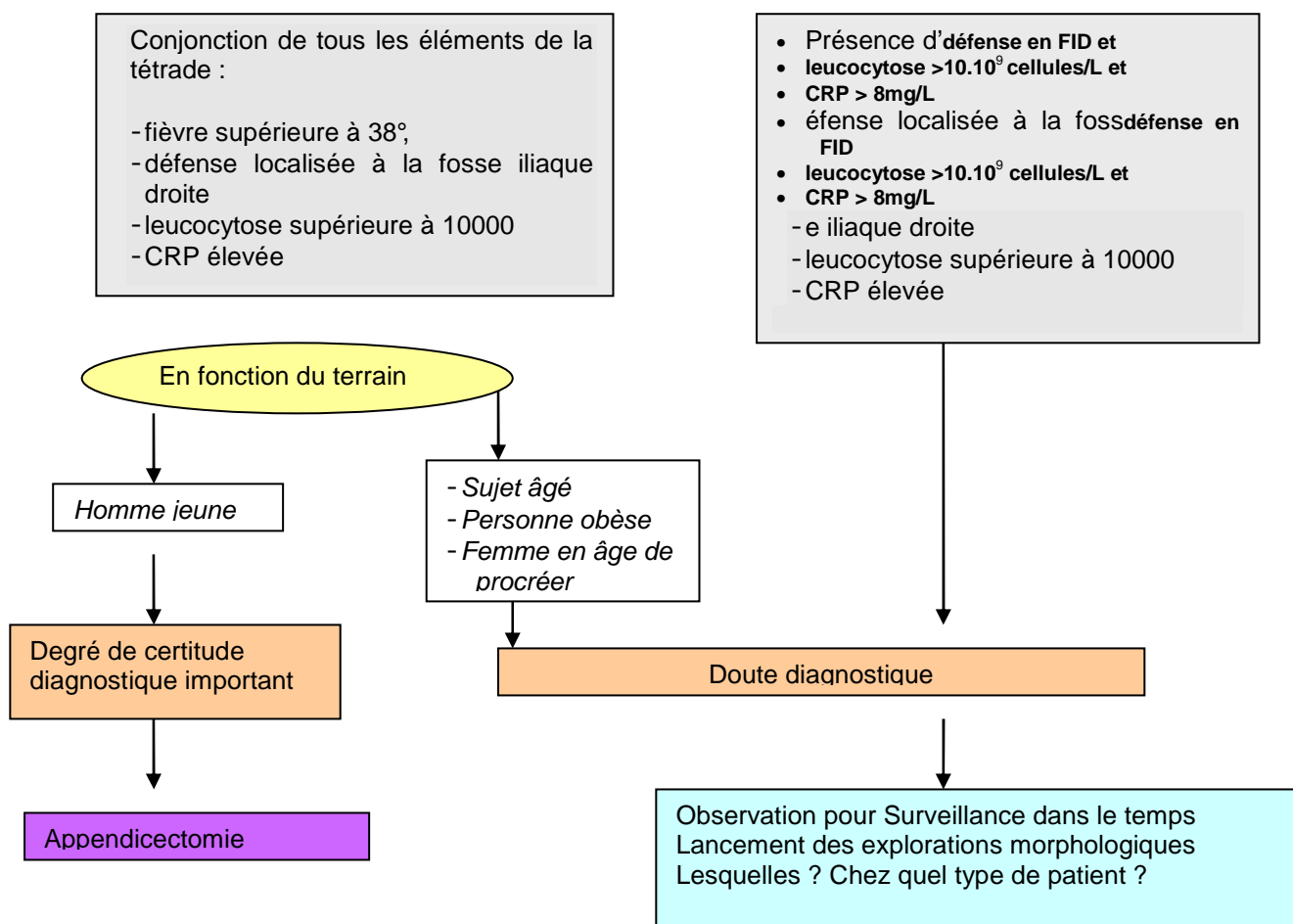
Concernant la comparaison de la performance diagnostique de l'échographie et du scanner, le groupe d'orientation estime que ces deux méthodes ne sont pas concurrentes, La fiabilité de l'une ou l'autre des deux méthodes dépend beaucoup du terrain et il faudrait tout au plus définir la place de chacune dans la stratégie diagnostique.

Concernant l'IRM, cette méthode d'imagerie est utilisée aujourd'hui pour un terrain particulier, celui de la femme enceinte. Compte tenu du manque de données robustes relatives aux autres situations cliniques, il ne paraît pas opportun d'intégrer cet examen dans la démarche diagnostique pour la population générale.

Enfin, le groupe d'orientation estime que l'interrogation d'un radiologue pédiatre serait utile pour savoir si les explorations chez l'enfant et le petit enfant nécessite une approche particulière.

« Que fait-on et quand ? »

Une première réflexion sur la construction d'un arbre décisionnel chronologique a été proposée, elle se schématise de la façon suivante :



En conclusion

Sur le plan méthodologique, cette première réunion du groupe a permis de dégager les éléments d'orientation suivants :

- proposition de préciser l'histoire naturelle de la maladie permettant de mieux introduire la notion d'appendicites ;
- nécessité d'élargir l'analyse de la littérature aux articles originaux ;
- la nécessité de compléter la composition qualitative du groupe d'orientation par un anatomopathologiste et par un radio-pédiatre (l'enfant de moins de 15 ans est une situation qui requiert des compétences spécifiques).

Après discussion générale, il a été décidé de délimiter le champ d'évaluation aux situations les plus courantes, à savoir :

- l'homme jeune ;
- la femme en âge de procréer ;
- le sujet âgé ;
- le cas de la femme enceinte et de l'enfant de moins de 15 ans nécessitent des approches spécifiques qui pourraient être envisagées dans une étape ultérieure.

En définitive, il s'agira dans ce travail de :

- définir la performance diagnostique des signes et symptômes cliniques et biologiques et plus particulièrement celle de la CRP ;
- définir la validité des scores clinico-biologiques ;
- définir la place de l'échographie et du scanner dans le processus diagnostique ;
- définir la place de la coelioscopie dans le processus diagnostique ;
- élaborer un arbre décisionnel qui tient compte du tableau clinique et du terrain.

Annexe 6. Compte rendu de l'audition du président de la SFMU du 06 janvier 2012

APPENDICECTOMIE

Éléments décisionnels pour une indication pertinente

Compte rendu de l'audition du Pr Jeannot Schmidt président de la SFMU (06 janvier 2012)

Représentants de la HAS /SEAP :

Dr Nadia Zeghari-Squalli PhD, Dr Michèle Morin Surroca, Dr Sun Hae Lee Robin

Point de vue de l'urgentiste

Observations sur les l'évolution du nombre d'appendicectomie sur les dix dernières années.

Les dernières années ont connu une diminution nette et progressive du nombre global d'appendicectomies sur le territoire national, elle serait imputable à :

- l'avènement des outils d'imagerie diagnostique ;
- un meilleur accès à l'imagerie ;
- une évolution du savoir faire des cliniciens et radiologues français

Dans ce contexte, le constat d'augmentation des situations d'appendicectomies compliquées peut sembler paradoxal. La raison principale revient à une prise en charge retardée dont l'origine peut être clinique et/ou organisationnelle :

- le patient se présente à un stade précoce avec un tableau clinique incomplet ;
- le patient arrive tôt mais avec un traitement antalgique ou anti-inflammatoire susceptible de masquer la symptomatologie, le patient est alors renvoyé et il revient quand la situation a évolué vers une complication ;
- l'examen d'imagerie réalisé trop tôt pour que la morphologie appendiculaire soit affectée.
- le manque d'expérience de certains radiologues pour qui la sémiologie d'un appendice malade n'est pas encore acquise ;
- dans certaines régions comme il est le cas à Clermont Ferrand, où les plateaux d'Urgence et de chirurgie digestive se situent sur deux sites différents, le transfert des patients d'un site à l'autre rallonge les délais de prise en charge post-diagnostic.

La dimension clinique

La seule présence d'une douleur abdominale, doit alerter sur le risque d'une appendicite. C'est invariablement le symptôme initial et il ne faut pas se limiter à la douleur en fosse iliaque droite typique. Peu de patients arrivent avec une forme complète du tableau typique, d'autres formes sont associées à d'autres signes.

Ainsi, devant une douleur abdominale :

- il faut prendre la température avec toujours une réflexion sur la prise d'antalgiques ou d'anti-inflammatoires par le patient ;
- si la douleur est typique et qu'elle n'est pas confrontée à d'autres éléments diagnostiques, il faut réaliser une imagerie ;
- l'échographie est à privilégier en première intention notamment chez la femme en âge de procréer et chez l'enfant ;

- le scanner est à privilégier chez la personne âgée en raison de la diversité des diagnostics différentiels, de même chez le sujet obèse si l'échographie s'avère non contributive.

La diversité de situations cliniques et l'absence dans la littérature de recommandations claires qui décrivent tous les signes et symptômes cliniques relatifs à une appendicite aiguë, amènent à penser qu'il serait intéressant, sur le plan national, de promouvoir une étude observationnelle dans le cadre d'un PHRC.

La dimension organisationnelle

Rôle de l'urgentiste et des différents intervenants

L'appendicite aiguë est une « vraie » urgence, le rôle de l'urgentiste dans cette situation est d'établir le diagnostic.

Le rôle du radiologue est de faire la discrimination diagnostique en cas de doute. Il doit être hors contexte global et ignorer les résultats clinico-biologiques pour ne pas être influencé dans son analyse.

Notion de parcours clinique

Le processus de prise en charge du patient sur le territoire national doit être harmonisé,

- il n'est pas recommandé un passage direct du patient du généraliste vers le chirurgien, cette démarche peut contribuer à augmenter le chiffre des appendicectomies (c'est une réalité qui est constatée sur le terrain) ;
- l'accès à l'examen d'imagerie doit être uniforme, l'existence d'un service de radiologie sur le site ne signifie pas une disponibilité absolue du radiologue. Par ailleurs, il est important de mettre en place un protocole national qui définit la technique d'analyse tomographique (en ce qui concerne les coupes) et les modalités d'utilisation des produits de contraste.

Aspect Indicateurs

Pour l'analyse rétrospective de la pertinence des appendicectomies réalisées, sur le territoire national, le collège français de médecine d'urgence pourrait, en partenariat avec la société française de radiologie, initier une étude prospective transversale, répétée à plusieurs reprises, avec une approche méthodologique réaliste privilégiant le choix de journées de type standard.

Il pourrait être considéré comme « faute de parcours de soins » le fait qu'un malade se fasse opérer là où il n'y pas d'imagerie.

Dans cette même démarche, le rôle de l'anatomopathologiste est prépondérant, il est nécessaire que les comptes-rendus soient clairs et précis pour permettre un contrôle de qualité a posteriori. Il n'est pas admis aujourd'hui de se cacher derrière des termes dont la signification est confuse « ex : appendicite subaiguë ».

Le taux acceptable d'appendicites blanches peut raisonnablement se situer autour de 15%.

A la question : *l'élaboration d'un référentiel peut-elle contribuer à diminuer le nombre d'appendicectomies ?*

La réponse est oui, une recommandation représente une contrainte, donc indéniablement influence le comportement. Cependant, en parallèle, il serait utile de mener une campagne d'information « Grand Public » qui aurait pour avantage supplémentaire de sensibiliser tous les acteurs impliqués dans le parcours de soins.

Annexe 7. Compte rendu de la deuxième réunion du groupe d'orientation du 03 mai 2012

APPENDICECTOMIE

Éléments décisionnels pour une indication pertinente

Compte rendu de la deuxième réunion du groupe d'orientation

03 mai 2012

Membres du groupe d'orientation

Membres présents : Dr Hervé BERCHE, Dr Stéphanie FRANCHI-ABELLA, Dr Philippe LEVEAU, Pr Philippe MARTEAU, Pr Christophe PONCELET, Pr Karem SLIM, Dr Nathalie SIAUVE, Pr Bertrand TOUSSAINT, Pr Benoit TERRIS,

Représentants de la HAS /SEAP :

Dr Nadia Zeghari-Squalli PhD, Dr Michèle Morin-Surroca, Dr Sun Hae Lee Robin,

Objet : Formalisation des conclusions issues de l'analyse de la littérature sous forme de «référentiels» et élaboration d'une stratégie de prise en charge diagnostique.

Lors de la réunion, l'analyse et la synthèse des données de la littérature ont été présentées aux membres du groupe d'orientation qui ont constaté la faiblesse des études disponibles.

A l'issue d'une discussion sur la base de ces éléments, les experts ont contribué à la finalisation des conclusions portant sur les différents éléments du diagnostic de l'appendicite, à l'élaboration de propositions de référentiels et à la construction d'un arbre décisionnel.

En préambule, les experts ont tenu à souligner que l'enjeu qui sous-tend ce travail est l'adoption d'une attitude diagnostique appropriée qui permette aussi bien de diminuer le risque des appendicites perforées que de réduire le nombre d'appendicectomies inappropriées.

Il a par ailleurs été rappelé la difficulté d'élaborer des propositions de référentiels qui soient claires dans un contexte où il existe une hétérogénéité des données recueillies, un faible niveau de preuves, un manque d'engouement pour cette « vieille affection », une absence de données scientifiques nouvelles et un indicateur de référence discuté (analyse histopathologique)

Les experts ont souligné, à ce propos, l'intérêt de disposer d'une définition histopathologique ad-hoc d'une appendicite aiguë avec un seuil à partir duquel on peut considérer qu'un appendice est pathologique. Les experts rapportent que parfois le seul traumatisme chirurgical peut induire un afflux de polynucléaires et suggérer une inflammation. Vraisemblablement, quand l'appendicite n'a pas une présentation macroscopique évidente (appendicites suppurées...), une variabilité inter-observateur parmi les pathologistes pourrait exister dans le diagnostic d'une appendicite débutante.

Éléments de diagnostic clinique et biologique

A la lumière des éléments recueillis dans la littérature et notamment la valeur diagnostique de la fièvre dans un tableau clinique de suspicion d'appendicite, les experts considèrent que cet élément n'est pas suffisamment discriminant pour compter parmi les signes cliniques décisionnels.

Concernant les éléments cliniques et biologiques qui permettent de poser le diagnostic d'appendicite, les membres du groupe d'orientation considèrent, sur la base des données de la littérature qu'aucun signe clinique ou biologique considéré individuellement n'a de valeur suffi-

samment discriminante pour permettre de poser le diagnostic d'appendicite. En revanche, la valeur diagnostique de la combinaison de certains signes cliniques et/ou biologiques permet de renseigner sur la probabilité diagnostique d'une appendicite.

Les membres du groupe d'orientation ont retenu les propositions suivantes :

Proposition 1

Devant une douleur abdominale, aucun signe ou symptôme clinique (y compris la fièvre ou l'apyrexie) n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.

Proposition 2

Devant une douleur abdominale, aucun paramètre biologique (leucocytes, granulocytes, PNN, CRP) n'a de valeur diagnostique individuelle suffisamment élevée pour permettre de confirmer ou d'exclure un diagnostic d'appendicite.

Proposition 3

Devant une douleur abdominale évocatrice d'une appendicite, l'absence concomitante d'une défense en fosse iliaque droite, d'une leucocytose supérieure à 10×10^9 cellules/L et d'une CRP supérieure à 8 mg/L rend le diagnostic d'appendicite aiguë peu probable.

Proposition 4

Devant une douleur abdominale évocatrice d'une appendicite, la combinaison d'une défense en fosse iliaque droite à une leucocytose supérieure à 10×10^9 cellules/L et une CRP supérieure à 8 mg/L chez l'homme jeune rend le diagnostic d'appendicite aiguë très probable.

► Les scores clinico-biologiques

Concernant les scores clinico-biologiques, les experts soulignent que l'utilisation de ces outils en France est très « service » et école de chirurgie dépendants.

Le score le plus connu et le plus décrit dans la littérature est le score d'Alvarado. Les données de la littérature ne permettent pas de valider de façon ferme ce score dont le principal biais est qu'il est construit à partir de données rétrospectives recueillies sur des patients déjà hospitalisés pour appendicite.

Des experts du GO citent un score français « le score de François » qui semble présenter des résultats satisfaisants, validée par une étude prospective publiée en 1991 (97). Cette étude sera analysée à la lumière des critères de sélection de la littérature adoptés pour ce travail.

Les experts ne recommandent pas l'utilisation des scores dans la pratique clinique mais considèrent néanmoins que ceux-ci pourraient aider à pondérer les évaluations cliniques des praticiens en termes de probabilité et constituer un outil pour la stratification du risque et l'orientation des patients vers des examens complémentaires.

Proposition 5

Au regard de la faible qualité des données publiées, les scores clinico-biologiques d'Alvarado et d'Andersson ne peuvent pas être considérés validés, mais peuvent apporter une aide pour apprécier la probabilité d'appendicite et pour orienter la suite de la prise en charge du patient.

Les examens d'imagerie

L'analyse de la littérature montre qu'il existe une grande hétérogénéité entre les techniques d'échographie et de scanner utilisées dans les différentes études publiées. Un expert du GO estime que le type d'échographe utilisé ne gêne pas la détection des appendicites. Un autre expert

nuance cette affirmation en considérant que cela pouvait être vrai il y a une dizaine d'années compte tenu des techniques disponibles alors, mais qu'avec l'avènement de technologies plus sophistiquées, cela l'est beaucoup moins aujourd'hui.

► L'échographie

Pour l'échographie, les experts radiologues du groupe jugent nécessaire de disposer des éléments cliniques avant la réalisation de l'examen, l'appréhension du contexte clinique permettant d'orienter le processus exploratoire et d'améliorer la performance diagnostique de cet examen.

Les membres du GO estiment que l'échographie a une bonne valeur diagnostique positive quand l'appendice est visualisé. Ils confirment cependant le caractère opérateur-dépendant « connu » de cet examen. Celui-ci est susceptible d'affecter la valeur de la performance diagnostique, en particulier pour les données recueillies dans les études américaines où les examens sont effectués par les techniciens et non par les radiologues. Les experts soulignent qu'il existe une courbe d'apprentissage qui rend les performances diagnostiques de l'échographie variables en fonction de l'expérience de l'opérateur d'où la variabilité inter-observateur. La situation est différente en Europe, cet examen est réalisé par le radiologue qui de part sa formation de clinicien peut prendre en compte des signes indirects (lame d'épanchement péricæcal ... etc) qui pourraient contribuer à la performance globale de cet examen.

Proposition 6 : « Valeur diagnostique de l'échographie »

D'après une littérature hétérogène et de faible qualité, l'intérêt diagnostique de l'échographie paraît variable, opérateur et patient dépendant. Lorsque l'échographie visualise l'appendice, sa valeur diagnostique semble être importante. En revanche, la non visualisation de l'appendice ne permet pas d'exclure une appendicite.

► Le scanner

Dans la littérature analysée, les études originales des méta-analyses présentent de nombreux biais susceptibles d'induire une surestimation des valeurs de la performance diagnostique du scanner.

Une hétérogénéité est, par ailleurs, relevée dans les études en ce qui concerne la présentation clinique, le degré de sévérité du tableau clinique, la variabilité des techniques de scanner, les modalités de contraste et les sites d'examen.

Selon les experts du GO, l'hétérogénéité des techniques de scanner de même que les modalités de contraste n'influent pas sur la performance diagnostique du scanner. Ce constat est confirmé dans la littérature qui montre que quelles que soient les techniques et les modalités, les données restent homogènes

Les experts rapportent que dans la pratique française, les modalités de réalisation du scanner sont variables d'un service à l'autre et parfois même au sein d'un même service. Cette variabilité tient du fait que la décision d'injection d'un produit de contraste par le radiologue dépend du contexte clinique individuel et en particulier de la présence ou pas de graisse intra-abdominale. Cet élément conditionne la qualité de l'examen et son interprétation. Le radiologue est plus à même d'utiliser un produit de contraste chez un patient mince que chez un patient obèse chez qui l'inflammation de la graisse est bien visualisée.

Sur ce point, les données de la littérature ne permettent pas, de trancher sur l'utilisation ou non d'un produit de contraste.

Les experts tiennent tout particulièrement à rappeler que l'imagerie scanner est un examen irradiant et alertent sur le risque du cumul des doses.

Les doses retrouvées dans la littérature et notamment dans la recommandation de l'ACR (30 mSv) sont discutés par les membres du groupe qui considèrent que ces valeurs d'irradiation ne correspondent pas au contexte français, les doses moyennes habituellement utilisées en France sont de

l'ordre de 12 mSv (sources IRSN et SFR) et selon l'un des radiologues du groupe les doses seraient plutôt actuellement autour de 3 mSv.

Lors de la réunion, il a par ailleurs été évoquée la publication récente d'une étude randomisée monocentrique coréenne publiée dans « The New England Journal of Medicine » qui compare un scanner à "faibles doses" (2 mSv environ, 116 mGy.cm de valeur médiane) et un scanner standard (8 mSv environ, 521 mGy.cm de valeur médiane) (91). Cette étude conclut au regard du taux d'appendicectomies négatives que le scanner à faible dose est non inférieur à un scanner avec une dose standard chez les jeunes adultes avec une suspicion d'appendicite.

Les experts rappellent néanmoins que la décision d'utiliser une dose « normale » ou une dose « faible » est prise en fonction de la morphologie du patient, dans le cas d'un patient mince n'ayant pas beaucoup de graisse, il est préférable d'avoir d'emblée une dose « normale » plutôt qu'une dose faible (afin d'éviter le cumul des doses si le premier scanner n'est pas conclusif). L'utilisation des scanners à faible dose ne peut par conséquent pas être généralisée.

Concernant le délai de réalisation du scanner en cas de suspicion d'appendicite, les experts se sont référés à un document de consensus, élaboré en 2004 par la Collégiale des radiologues, la Collégiale des urgentistes et la Direction de la politique médicale de l'AP-HP (92).

Ce document préconise la réalisation d'un scanner dans les 4h en cas de signes cliniques de gravité (masse à la palpation, empatement...) ou de terrain fragile (corticothérapie, diabète...) et dans un délai de 24 h en absence de signes cliniques de gravité.

Après discussion des données de la littérature et à la lumière des éléments de la pratique française, les membres du GO concluent :

Proposition 7

Malgré les biais présents dans les études, les résultats sont concordants et montrent que le scanner est performant pour établir le diagnostic d'appendicite ou pour l'exclure. L'indication de cet examen doit être posée en considérant son caractère irradiant.

► Comparaison scanner vs échographie

Dans des études comparant le scanner et l'échographie chez le même patient, le scanner a une performance diagnostique supérieure à celle de l'échographie. Toutefois, ces études qui définissent la performance diagnostique ne permettent pas de définir la place respective de chacun de ces examens dans la démarche diagnostique, celle-ci devant tenir compte des caractéristiques individuelles du patient.

Les membres du GO considèrent qu'il ne faut pas opposer l'échographie au scanner, une telle démarche serait déconnectée de la réalité du terrain où il est considéré que pour certains groupes de patients (femme jeune et patient mince), il est plus pertinent de proposer l'échographie en premier, tandis que pour d'autres, le scanner peut être proposé d'emblée (personnes âgées et personnes obèses).

Par conséquent, les experts considèrent qu'il y a une place pour chacune des deux techniques, celle-ci dépend du contexte clinique qui définit l'ordre chronologique décisionnel.

Les membres du Go concluent

Proposition 8

Le scanner et l'échographie ne sont pas des examens concurrents, ils peuvent avoir chacun leur place dans la démarche diagnostique.

► La radio d'abdomen sans préparation (ASP)

Les membres du groupe de travail souhaitent qu'il soit indiqué clairement dans le rapport final, que conformément aux recommandations de la HAS (26), la radio d'abdomen sans préparation (ASP) n'est pas indiquée pour le diagnostic de l'appendicite.

► L'IRM

Concernant l'IRM, cette méthode d'imagerie est utilisée aujourd'hui pour un terrain particulier, celui de la femme enceinte et de l'enfant. Dans le champ retenu pour cette évaluation qui exclue ces deux populations, il n'a pas été identifié de données probantes dans la littérature. De plus, compte tenu du parc IRM français et notamment du déficit en matière d'équipements, il ne paraît pas opportun à ce jour (et avant amélioration de la situation) d'intégrer cet examen dans la démarche diagnostique d'appendicite pour les autres populations.

La coéloscopie exploratrice

Au regard des éléments de la littérature analysée, les membres du groupe de travail considèrent que la coéloscopie exploratrice n'est pas un examen de première intention dans un contexte de suspicion d'appendicite, elle peut être proposée lorsque l'imagerie (échographie et scanner) n'est pas contributive et qu'il persiste un doute notamment chez la femme jeune.

Les experts rappellent que la coéloscopie diagnostique est un examen invasif qui présente un taux de morbidité de l'ordre de 2% (98).

Proposition 9 :

Il n'y a pas de données de qualité suffisante dans la littérature pour positionner la coéloscopie exploratrice dans la démarche diagnostique.

La coéloscopie exploratrice n'est pas à réaliser en première intention, mais elle peut être proposée lorsque l'imagerie (échographie puis scanner) n'est pas contributive chez la femme jeune.

Arbre décisionnel devant une douleur abdominale chez l'adulte

Les experts considèrent qu'avant de construire un arbre décisionnel avec une démarche diagnostique, il convient, devant une douleur abdominale, de définir 3 situations cliniques avec des niveaux de probabilités différents : fort, faible et intermédiaire. Ces situations tiennent compte de l'âge, du sexe et du morphotype des patients.

Il est important de préciser que cet arbre décisionnel s'adresse à l'adulte, il ne concerne pas les situations particulières que sont l'appendicite de l'enfant et de la femme enceinte. Il tient compte en particulier des situations de forte prévalence notamment les sujets de moins de 30 ans et les plus de 65 ans.

Les situations de faible probabilité diagnostique

Devant une douleur abdominale, les situations considérées comme présentant une faible probabilité d'appendicite sont celle où il y a une absence concomitante de la triade suivante :

- Absence de défense abdominale,
- Absence de leucocytose supérieure à 10.10^9 ;
- Absence de CRP supérieure à 8mg/L.

Dans cette situation la conduite à tenir est la surveillance avec une modalité qui devra être adaptée au contexte médical, familial et social (ex : pour les gens du voyage, il est plus prudent de les garder en observation pour éviter une aggravation éventuelle du tableau).

Dans cette situation la conduite à tenir est la surveillance avec une modalité qui devra être adaptée au contexte médical, familial et social (ex : nomades, sans domicile fixe, phobie du milieu

médical,...etc.) il est plus prudent de les garder en observation pour éviter une aggravation éventuelle du tableau

Les situations de forte probabilité diagnostique

Une seule situation apparaît comme présentant une forte probabilité d'appendicite en présence d'une douleur abdominale, il s'agit de l'association concomitante chez l'homme jeune (âge \leq 30 ans) des trois éléments suivants :

- une défense en FID ;
- une leucocytose $>10.10^9$ cellules/L ;
- une concentration de CRP supérieure à 8mg/L.

Une deuxième situation, également décrite chez l'homme jeune et reposant uniquement sur des signes cliniques est décrite dans la littérature. Il s'agit de la présence concomitante de 3 signes cliniques suivants : une défense en FID, une douleur migrante et une contracture abdominale. Ce tableau très rare par ailleurs (moins de 3% des cas) est selon les experts plus évocateur d'une péritonite.

Les situations de probabilité diagnostique intermédiaire ou « doute »

En dehors des situations où les 3 éléments de la triade sont absents (faible probabilité) et la situation où ils sont présents de façon concomitante chez l'homme jeune (forte probabilité), toutes les autres situations relèvent du doute diagnostique et présentent une probabilité intermédiaire. Ces situations nécessitent des examens complémentaires d'exploration morphologique (échographie ou scanner). Le recours à l'une ou l'autre des deux techniques d'imagerie dépend du terrain.

Chez la femme jeune, il faut privilégier l'échographie en première intention afin d'éviter l'irradiation et d'exclure une pathologie gynécologique. Si l'échographie n'est pas contributive, il faut faire un scanner et en cas de persistance de doute, il faudrait s'orienter vers un avis gynécologique.

Chez le sujet âgé, il faut privilégier le scanner, la recherche s'oriente d'emblée vers un diagnostic différentiel colique. En cas de scanner négatif il faudrait garder le patient en observation pendant 24h ou l'orienter vers une prise en charge ad-hoc si un autre diagnostic est avéré.

Chez le sujet obèse, la décision de réaliser le scanner ou l'échographie en première intention reste à la discrétion du radiologue. Il a été, toutefois, rappelé les limites de la performance diagnostique de l'échographie dans cette situation.

Si l'échographie réalisée en première intention n'est pas conclusive, il faudrait réaliser un scanner. Si ce dernier est négatif, il faudrait garder le patient en observation ou l'orienter vers la prise en charge ad-hoc si une autre pathologie est détectée.

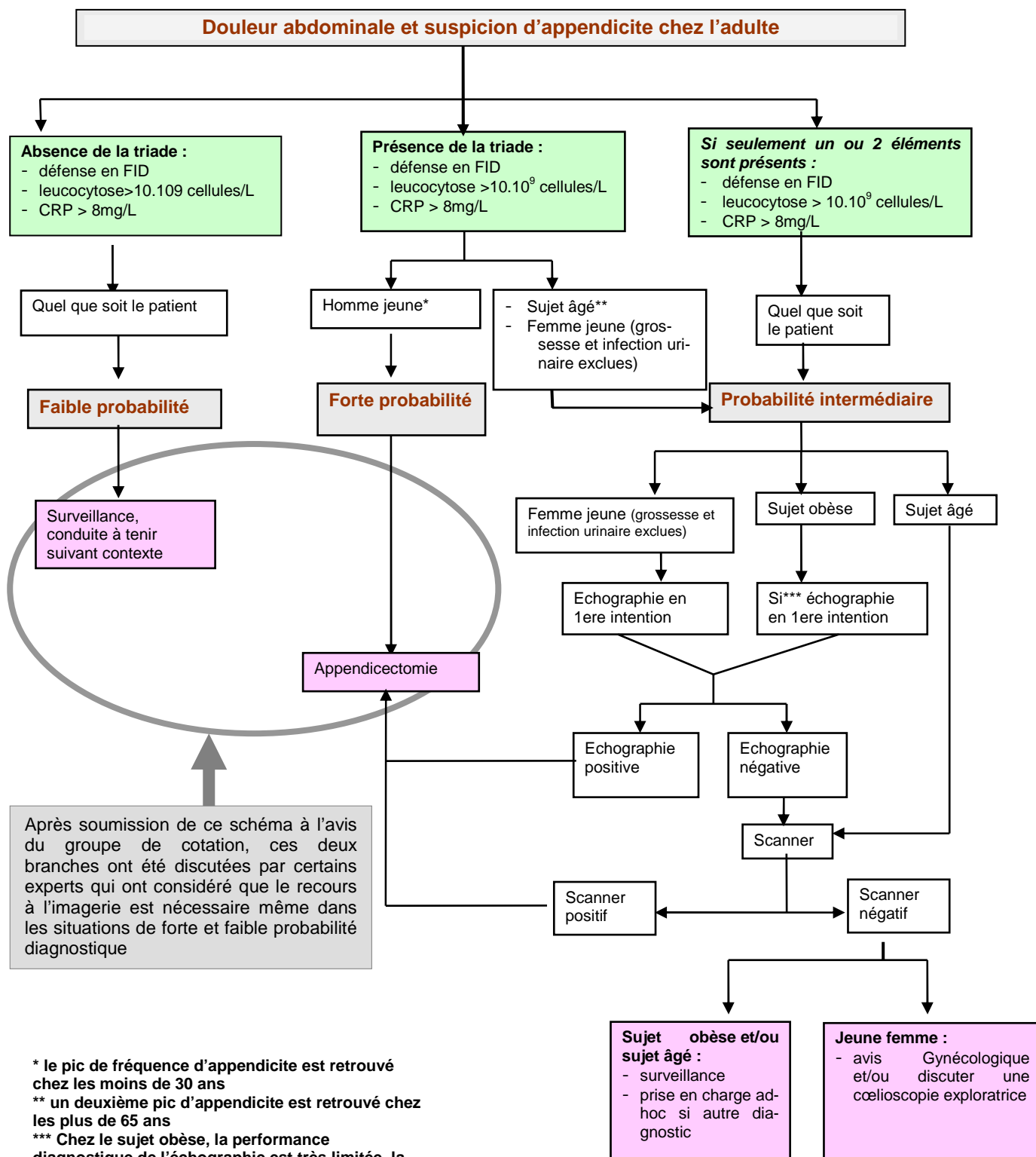
Le délai retenu pour la réalisation des examens d'imagerie est de 4h en cas de signes cliniques de gravité et de 24 h en absence de signes cliniques de gravité.

Partant des ces différents éléments, un arbre décisionnel a été construit en tenant compte des ces différentes situations cliniques.

Il convient cependant de rappeler que ce chemin est possible quand l'établissement de santé dispose d'un environnement adéquat : plateau technique et professionnels de santé. Or, de l'avis des membres du groupe d'orientation, cela ne semble pas être toujours le cas.

Dans la pratique, ces schémas s'adaptent aux conditions locales et la problématique organisationnelle peut contraindre à adopter une orientation plutôt qu'une autre.

Proposition d'arbre décisionnel devant une suspicion d'appendicite soumise au groupe de cotation



* le pic de fréquence d'appendicite est retrouvé chez les moins de 30 ans
 ** un deuxième pic d'appendicite est retrouvé chez les plus de 65 ans
 *** Chez le sujet obèse, la performance diagnostique de l'échographie est très limitée, la décision de réaliser le scanner ou l'échographie en 1^{er} intention est à la discrétion du radiologue

Annexe 8. Niveau de preuve scientifique fourni par le grade des recommandations

Définition du grade des recommandations

Niveau de preuve scientifique fourni par la littérature	Grade des recommandations
Niveau 1 <ul style="list-style-type: none"> essai comparatif randomisé de forte puissance méta - analyse d'essais comparatifs randomisés analyse de décision basée sur des études bien menées 	A Preuve scientifique établie
Niveau 2 <ul style="list-style-type: none"> essai comparatif randomisé de faible puissance étude comparative non randomisée bien menée étude de cohorte 	B Présomption scientifique
Niveau 3 <ul style="list-style-type: none"> étude cas - témoins 	C
Niveau 4 <ul style="list-style-type: none"> étude comparative comportant des biais importants étude rétrospective série de cas 	Faible niveau de preuve

Annexe 9. Tableaux d'évidence

Tableau 14. Méta-analyses sur les symptômes et signes cliniques R.E.B Anderson, 2004 (51)

Auteur, année, référence	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Test de référence	Résultats Et conclusion	Limites méthodologiques
<p>Anderson, 2004 (51)</p> <p>Méta-analyse</p>	<p>Analyser dans les situations de suspicion d'appendicite la valeur diagnostique de :</p> <p>1- l'histoire de la maladie, des signes cliniques</p> <p>2-des variables biologiques</p>	<p>Période: 1994-2004</p> <p>Bases : Medline</p> <p>Mots clés renseignés</p> <p>Langue : anglais, allemand, français, italien, espagnol, portugais et langues scandinaves</p>	<p>Articles sélectionnés : patients admis pour une suspicion d'appendicite, données d'efficacité diagnostique (RVP, RVN, ROC)</p> <p>24 prospectives consécutives</p> <p>Articles exclus, population pédiatrique difficulté d'extraire les données résultats incohérents</p>	<p>Diagnostic histopathologique pour les patients opérés (23 études)</p> <p>Pour les patients non opérés, absence de survenue d'appendicite pendant le suivi (5 études)</p>	<p>taux médian d'appendicites 41 % (26,7 à 60,6 %). taux médian d'appendicites perforées 17,4% (3,7 à 28,6%) taux médian d'appendicites négatives 20,1% (5,2-42,2%).</p> <p>I-Histoire de la maladie, signes et symptômes <i>courbes ROC, (IC 95%) p<0,001</i></p> <p>Douleur migrante : 0,68 (0,63-0,74) Défense : 0,68 (0,60-0,76) Indirect tenderness : 0,65 (0,55-0,75) Douleur à la décompression : 0,70 (0,65-0,75) <i>RVP p<0,001</i></p> <p>Douleur migrante : 2,06 (1,63-2,60) Défense : 2,48 (1,60-3,84) Sensibilité (douleur) indirecte : 2,47 (1,38-4,43) Douleur à la décompression : 1,99 (1,62-2,45) <i>RVN</i></p> <p>Douleur migrante : 0,52 (0,40-0,69) p<0,001 Défense : 0,57 (0,48-0,68) p=0,015 Sensibilité indirecte : 0,71 (0,65-0,77) p=0,082 Douleur à la décompression : 0,39 (0,32-0,48) p=0,004</p> <p>II-Variables biologiques <i>courbes ROC, (IC 95%)</i></p> <p>Leucocytose : 0,77 (0,75-0,78) p=0,171</p>	<p>L'absence de description des données démographiques âge et sex-ratio</p> <p>Population très hétérogène (variabilité des taux d'appendicites + et – et du taux d'appendicites perforées]</p> <p>Le test de référence n'est pas appliqué à tous les patients</p> <p>Seules 5 études sur 24 rapportent les résultats de suivi des patients non opérés</p> <p>Pour les variables individuelles, les paramètres diagnostiques ne sont pas tous calculés sur l'ensemble de la population de la méta-analyse.</p> <p>Pour les variables</p>

				<p>Taux de Granulocytes 0,78 (0,75-0,82] p=0,370 Taux de PMN : 0,77 (0,70-0,84] p<0,001 Taux de CRP : 0,75 (0,68-0,85] p<0,001</p> <p><i>RVP</i></p> <p>Leucocytose $\geq 10^{10}/l$: 2,47 à 3,47 Taux de Granulocytes $\geq 7 \cdot 10^9/l$: de 1,64 à 7,09 Taux de PMN > 85 % : 3,82 (2,86-5,08] Taux de CRP > 20 mg/l : 2,39 (1,67-3,41]</p> <p><i>RVN</i></p> <p>Leucocytose $\geq 10^{10}/l$: 0,26-0,81 Taux de Granulocytes $\geq 7 \cdot 10^9/l$: de 0,31-0,74 Taux de PMN > 85 % : 0,58 (0,51-0,66] Taux de CRP > 20 mg/l : 0,47 (0,28-0,81]</p> <p>Fièvre > 37,7 VPP=1,57 VPN=0,65 > 38,5 VPP=1,87 VPN=0,89</p> <p>Faible pouvoir discriminant des variables individuelles</p> <p>Les signes d'irritation péritonéale, la migration de la douleur et les signes biologiques de l'inflammation (ont un haut pouvoir discriminant et devraient être inclus dans la stratégie diagnostique des patients avec une suspicion d'appendicite</p> <p>III- Combinaison des variables <i>courbes ROC, (IC 95%)</i></p> <p>GB > 10×10^9 + CRP > 8mg/L AUC ROC = 0,96 (0,92-1,00] GB > 10×10^9 + CRP > 12 mg/L + PNN > 70% AUC ROC = 0,79 (0,74-0,84] GB > 10×10^9 + CRP > 8mg/L + IL6 > 60ng/L AUC ROC = 0,87 (0,80-0,94] Défense + GB > 10×10^9 AUC ROC = 0,84 (0,80-0,88] GB > 10×10^9 + CRP > 12 mg/L AUC ROC = 0,85 (0,80-0,90]</p>	<p>combinées, la méta-analyse n'est pas réalisée. Simple description des résultats de 9 études originales de faible effectif (<60 patients) sans description de la qualité méthodologique des études</p>
--	--	--	--	---	---

					<p><i>RVP</i></p> <p>GB>10x10⁹ +CRP>8mg/L RVP= 23,32 (6,87-84,79]</p> <p>GB>10x10⁹ +CRP>12 mg/L+ PNN >70% RVP=20,85 (5,47-80,27]</p> <p>GB>10x10⁹ +CRP>8mg/L+IL6>60ng/L RVP=16,96 (3,08-98,66]</p> <p>Défense + GB>10x10⁹ RVP=11,34 (6,65-19,56]</p> <p>GB>10x10⁹ +CRP>12 mg/L RVP=8,22 (4,7-14,38]</p> <p><i>RVN</i></p> <p>GB>10x10⁹ +CRP>8mg/L RVN=0,03 (0,00-0,14]</p> <p>GB>10x10⁹ +CRP>12 mg/L+ PNN >70% RVN=0,03 (0,01-0,16]</p> <p>GB>10x10⁹ +CRP>8mg/L+IL6>60ng/L RVN=0,03 (0,01-0,16]</p> <p>Défense + GB>10x10⁹ RVN=0,14 (0,08-0,24]</p> <p>GB>10x10⁹ +CRP>12 mg/L RVN=0,05 (0,01-0,18]</p>	
--	--	--	--	--	--	--

Méta-analyse d'Anderson, 2004 (51)

Evaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori est fourni.
- La sélection des études de même que l'extraction des données n'a pas été faite par deux personnes
- La recherche documentaire n'est pas exhaustive.
- Il n'est pas précisé si le type de publication constitue un critère d'inclusion.
- La liste des études incluses et exclues n'est pas fournie.
- Les caractéristiques des études incluses ne sont pas indiquées.
- La qualité scientifique des études incluses n'a pas été évaluée.
- La qualité scientifique des études incluses dans la revue est utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites de façon sommaire.
- La probabilité d'un biais de publication n'a pas été évaluée.
- La présence ou l'absence de conflits d'intérêts n'ont pas été déclarées.

Seuls 2 critères de la grille AMSTAR sont remplis, la qualité de cette étude en est par conséquent jugée très faible.

Tableau 15. Revue systématique sur les symptômes et signes cliniques James M. Wagner 1996, (52)

Auteur, année, référence, pays	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Test de référence	Résultats Et conclusion	Limites méthodologiques
Wagner et al., 1996 (52)	Evaluation de la précision diagnostique des signes et symptômes cliniques	1966-1996 Stratégie de recherche : NR	1-Population adulte 2-Données permettant de calculer les rapports de vraisemblance Modalité d'analyse critique : NR Etudes analysées : 5 rétrospectives 7 prospectives	Confirmation anatomopathologique d'appendicite	<p><u>Caractéristiques des patients :</u></p> <ul style="list-style-type: none"> - Intervalles d'âge dans les études : borne inférieure de 4 à 15 ans et borne sup de 50 à 90 ans. - Pourcentage moyen de femmes 41% (entre 39 à 58% avec 1 étude 0% de femmes). - Présentation clinique de l'appendicite : suspicion d'appendicite (2886 patients), douleur abdominale aigue évoluant depuis moins d'une semaine (5275 patients) <p><u>Rapports de vraisemblances</u></p> <p>RVP</p> <ul style="list-style-type: none"> 1-Douleur en FID RVP=8 2-défense de l'abdomen RVP=4 3-douleur migrante de la région péri ombilicale à la FID=3,1 <p>RVN</p> <ul style="list-style-type: none"> 1- migration de la douleur (RVN=0,5, 2- douleur dans FID (RVN= 0,2) et 3- douleur similaire auparavant (RVN= 0,3) <p>A l'examen clinique, aucun élément ne permet à lui seul d'affirmer ou d'exclure le diagnostic d'appendicite</p>	<ul style="list-style-type: none"> - Prévalence non renseignée - Durée du suivi des patients non opérés : NR - Les paramètres diagnostiques ne sont pas calculés sur l'ensemble de la population de l'étude - Les études analysées sont toutes de niveau de preuves 2 et 3

Revue systématique de Wagner et al., 1996 (52)

Evaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori n'est pas fourni.
- La sélection des études de même que l'extraction des données n'a pas été faite par deux personnes
- La recherche documentaire n'est pas exhaustive.
- Il n'est pas précisé si le type de publication constitue un critère d'inclusion.
- La liste des études incluses et exclues n'est pas fournie.
- Les caractéristiques des études incluses ne sont pas indiquées.
- La qualité scientifique des études incluses n'a pas été évaluée.
- La qualité scientifique des études incluses dans la revue est utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites de façon sommaire.
- La probabilité d'un biais de publication n'a pas été évaluée.
- La présence ou l'absence de conflits d'intérêts n'ont pas été déclarées.

Aucun des critères de la grille AMSTAR n'est rempli, la qualité de cette étude en est par conséquent jugée très faible.

Tableau 16 Etude prospective de Laméris *et al.* (47)

Auteur, année, référence, Type d'étude	Objectif de l'étude	Population* Caractéristiques des patients	Critères de jugement « groupe d'experts »	Données recueillies	Résultats et conclusions	Limites méthodologiques
Laméris <i>et al</i> 2009 (47) Etude prospective consécutive multicentrique	Evaluer la performance diagnostique précision diagnostique des signes et symptômes cliniques et des variables biologiques dans le diagnostic de l'appendicite aiguë.	Patients adultes > 18 ans avec douleur abdominale aiguë évoluant depuis au moins 2 h et moins de 5 jours. 942 patients Age moyen 47 ans \pm 17,3 (19-94 ans) Sex-ratio F/H=1,22 (55% de femmes)	Diagnostic final réalisé chez tous les patients par un groupe d'expert sur la base des tous les éléments suivants : - Examen clinique - Numération sanguine - dosage de la CRP - examen anatomopathologique pour les patients opérés - Rapport d'imagerie - éléments de suivi clinique sur 6 mois	- Examen clinique - Numération sanguine totale, - Dosage de la CRP, - Âge, - Sexe, - Résultats d'imagerie : Radiographie de l'abdomen + échographie+ scanner, - éléments de suivi clinique sur 6 mois. <u>L'évaluation de la fréquence des différents profils clinico-biologiques</u> Les profils sont identifiés en fonction de : l'âge>30 ou < 50 ans, du sexe des signes biologiques (Leucocytes et CRP) et des signes cliniques de l'appendicite dite « classique » : (douleur migrante, sensibilité en FID, défense abdominale)	<u>Prévalence appendicite</u> : 30% <u>Variables individuelles</u> pour toutes les variables cliniques et biologiques AUC courbes ROC varie de 0,50 à 0,65. valeur prédictive la plus forte retrouvée pour : -sexe masculin (RVP=2) -leucocytes RVP= 2,1 pour un nombre compris entre 15.10 ⁹ et 20.10 ⁹ /L. <u>Combinaison des variables biologiques</u> hyperleucocytose >10.10 ⁹ et CRP>12mg/L, RVP= 1,4 (1,2-1,7] IC 95%. RVN=0,09 (0,03-0,3] IC 95%. <u>Fréquence des différents profils clinico-biologiques</u> -23 profils identifiés -5 profils présentent un taux d'appendicite >85% en particulier chez les hommes. -la fréquence relative de ces 5 profils dans la population de l'étude varie de 2 à 13 %. - le profil type « <i>appendicite classique</i> » (migration de la douleur en FID, + sensibilité en FID + défense), représente 6% de l'ensemble des patients. Dans ce groupe 100% d'appendicite chez les hommes et seulement 46 % chez les femmes. Chez les patients avec suspicion d'appendicite, le pouvoir discriminant (AUC) des données individuelles cliniques et biologiques est faible. La combinaison de caractéristiques cliniques et biologiques offrant une précision diagnostique élevée est relativement rare chez les patients avec	Les données cliniques des vrais-négatifs données ne sont pas incluses dans l'analyse (patients chez qui l'imagerie n'a pas été jugée nécessaire et n'a pas été réalisée)

				La précision diagnostique de la combinaison des différents éléments (profil) est exprimée par le taux d'appendicite au sein de chaque groupe clinico-biologique	suspicion d'appendicite.	
--	--	--	--	---	--------------------------	--

Tableau 17. Etude prospective de Cardall *et al* (54)

Auteur, année, référence, Type d'étude	Objectif de l'étude	Population* Caractéristiques des patients	Critères de jugement	Données recueillies	Résultats et conclusions	Limites méthodologiques
Cardall <i>et al</i> 2004 (54) Prospective consécutive monocentrique	Evaluer la capacité discriminante de la température et du nombre total de leucocytes chez les patients se présentant aux urgences avec suspicion d'appendicite	Patients présentant des signes et symptômes cliniques faisant suspecter une appendicite. 293 patients inclus dont 130 de sexe masculin et 163 de sexe féminin Age moyen 30.8 (7-75 ans)	Résultat histopathologique après chirurgie suivi téléphonique du patient pendant 2 semaines	Nombre de leucocytes Mesure de la température	Prévalence appendicite 31 % (92 patients) 21% des appendicites sont perforées Mesure du nombre total de leucocytes chez 274 patients. Leucocytes $>10. 10^9/L$: - Se 76 % (65%-84 %) IC 95%, - Sp 52% (45%-60%) IC 95% - RVP= 1,59 (1,31-1,93] IC 95% - RVN=0,46 (0,31-0,67] IC 95% - ROC, AUC=0,72 (0,65-0,79] IC 95%., seuil optimal, à $11,9 \times 10^9$ cellules/L avec Se=69% et Sp=72% Température mesurée chez 293 patients $>99F$ soit $37,22\text{ }^\circ\text{C}$ - Se=47% (36%-57%) IC 95% - Sp=64% (57%-71%) IC 95% - RVP= 1,3 (0,97-1,72] IC 95% - RVN= 0,82 (0,65-1,01] IC 95% - AUC ROC=0,59 (0,52-0,66] IC 95%. Conclusion Considérés chacun individuellement, le nombre total de leucocytes de même que la température corporelle ne permettent une discrimination entre les patients avec et sans appendicite	Biais de sélection (critères d'inclusion mal définis) Les cliniciens connaissent à priori les résultats (leucocytes et T [°]) avant de poser le diagnostic. Les rapports de vraisemblances sont calculés à partir de variables continues

Tableau 18. Etude prospective d'Ortega-Deballon *et al* (46)

Auteur, année, référence,	Objectif de l'étude	Population*	Critères de jugement	Données recueillies	Résultats et conclusions	Limites méthodologiques
Type d'étude		Caractéristiques des patients				
<p>Ortega-Deballon et al., 2008 (46)</p> <p>Prospective consecutive monocentrique</p>	<p>objectif principal : évaluer l'utilité des paramètres inflammatoires dans la gestion des adultes avec une douleur aiguë à la FID</p> <p>objectif secondaire : évaluer la corrélation entre les niveaux des marqueurs de l'inflammation et la gravité de l'appendicite</p>	<p><u>Critères d'inclusion :</u></p> <ul style="list-style-type: none"> - douleur abdominale aiguë en FID - âge > 15 ans <p><u>Critères d'exclusion</u></p> <ul style="list-style-type: none"> - femmes enceintes - appendicectomie antérieure - patients avec une pathologie ou traitement au long cours impactant la réponse inflammatoire. <p>134 patients Sex ratio F/H = 1,23 (60 hommes et 74 femmes) Age moyen 33 ans (31,7-34,3] IC 95% sur une population dont l'âge varie entre 15 et 75 ans</p>	<ul style="list-style-type: none"> - Résultat histopathologique - Examen macroscopique - Rapport d'imagerie - Observation sur 24 h et suivi téléphonique sur 72 h. 	<ul style="list-style-type: none"> - Age, sexe, - Histoire de la maladie - Température - Délai entre le début des signes et l'admission aux urgences - Dosage de la CRP (normalité <6mg/L) - Numération leucocytaire - Taux de granulocytes) - Examens histopathologiques pour les d'appendicectomies 	<p>Distribution du diagnostic en 2 sous groupes :</p> <p><u>Pathologies non-chirurgicales</u></p> <p>13,5% appendicectomies négatives 12,7% patients non opérés suivis</p> <p><u>Pathologies chirurgicales</u></p> <p>65,7% d'appendicites dont 17% perforées 8,2 % opérés pour autres pathologies chirurgicales</p> <p><u>1- Valeurs moyennes des marqueurs</u></p> <ul style="list-style-type: none"> - CRP= 67,7 mg/L (p<0,0005) - Leucocytes= 14846/mm³ (p<0,0005) - %granulocytes 79,9 % (p<0,001) <p><u>2-Capacité discriminante (AUC)</u></p> <ul style="list-style-type: none"> - CRP=0,846 (p<0,0005) - Leucocytes= 0,753 (p<0,0005) - % granulocytes= 0,685 (p<0,001) - Leucocytes +CRP=0,868 (p<0,0005) <p><u>Valeurs prédictives appendicite</u></p> <ul style="list-style-type: none"> - CRP Se=90,9% Sp=74,3% VPP=79,8% VPN= 74,3% - Leucocytes Se=86,4% Sp=42,9% VPP=73,6% VPN= 62,5% - % granulocytes Se=82,8% Sp=45,7% VPP=73,3% VPN= 48,5% - Leucocytes +CRP Se=85,8% Sp=37,1% VPP=93,2% VPN= 92,3% <p><u>CRP (mg/L) et sévérité de l'appendicite</u></p>	<p>Biais de sélection, l'étude s'adresse exclusivement aux patients présentant une douleur en FID.</p> <p>Prévalence d'appendicite élevée (65,7%) dans la population étudiée, augmentation de la performance des diagnostique des marqueurs étudiés</p>

					<ul style="list-style-type: none"> - Pas d'appendicite 12.2 (3.9-20.5] p<0,0005 - appendicite phlegmoneuse 20.1 27 (18.9-35.1] p<0,0005 - appendicite gangréneuse 60.6 (41.1-80] p<0,0005 - appendicite perforée 174.8 (107.8-241.8] p<0,0005 <p><u>Evolution en fonction du temps</u></p> <p>Délai >12h</p> <ul style="list-style-type: none"> - AUC ROC CRP=0,896 (p<0,0005] - AUC ROC Leucocytes=0,795 (p<0,0005] - AUC ROC (CRP+leucocytes)=0,913 (p<0,0005) <p>Délai<12h</p> <ul style="list-style-type: none"> - AUC ROC CRP=0,779 (p<0,0005) - AUC ROC Leucocytes=0,609 (p<0,0005) - AUC ROC (CRP+leucocytes)=0,845 (p<0,0005) <p>Conclusions</p> <p>La protéine C-réactive est un marqueur utile dans le traitement des patients avec une douleur en fosse iliaque droite. Elle augmente avec l'évolution du processus inflammatoire. Sa valeur prédictive est améliorée quand elle est combinée à la numération leucocytaire.</p> <p>Un patient avec protéine C-réactive et des leucocytes normaux a une probabilité très faible d'avoir une appendicite.</p>	
--	--	--	--	--	--	--

Tableau 19. Méta-analyse sur le score d'Alvaradon Robert Ohle *et al* 2011 (57)

Auteur, année, référence,	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Critères de jugement	Résultats Et conclusion	Limites méthodologiques
<p>Ohle <i>et al.</i>, 2011 (57)</p> <p>Revue systématique avec Méta-analyse</p>	<p>Evaluer la performance diagnostique du score ALVARADO</p>	<p>Période: janvier 1996-avril 2011</p> <p>Bases : EMBASE, Cochrane Library, MEDION et DARE</p> <p>Combinaison de mots clés et termes MeSH</p> <p>Inclusion d'articles cités, pas de restriction de langue</p>	<p>Critères de sélection :</p> <ul style="list-style-type: none"> - patients admis avec symptômes suggérant une appendicite (douleur abdominale, douleur en FID) -objectif : comparer le résultat du score avec l'examen histopathologique ou le suivi patient (chez les patients non opérés) -résultats présentés de manière à permettre l'extraction de données pour l'évaluation de la performance diagnostique en fonction des 3 différents niveaux de score Evaluation de la qualité méthodologique des articles Grille QUADAS Articles retenus : 31 études prospectives 10 études 	<ul style="list-style-type: none"> - Diagnostic histopathologique pour les patients opérés - Pour les patients non opérés : suivi 	<p>Précision diagnostique du score Alvarado</p> <p>Score 5 : admission pour observation</p> <ul style="list-style-type: none"> - sur l'ensemble des études Se= 0,99 (0,97-0,99] IC 95% - hommes Se=0,96 (0,88-0,99] IC 95% - femmes Se=0,99 (0,92-0,99] IC 95% - enfants Se=0,99 (0,83-1] IC 95% <p>- sur l'ensemble des études Sp=0,43 (0,36-0,51] IC 95%</p> <ul style="list-style-type: none"> - hommes Sp=0,34 (0,24-0,47] IC 95% - femmes Sp=0,35 (0,14-0,64] IC 95% - enfants Sp=0,57 (0,41-0,72] IC 95% <p>Score 7 : critère de chirurgie</p> <ul style="list-style-type: none"> - sur l'ensemble des études Se= 0,82 (0,76-0,86] IC 95% - hommes Se=0,88 (0,75-0,95] IC 95% - femmes Se=0,86 (0,78-0,92] IC 95% - enfants Se=0,87 (0,76-0,93] IC 95% <ul style="list-style-type: none"> - sur l'ensemble de la population Sp=0,81 (0,76-0,85] IC 95% - hommes Sp=0,57 (0,40-0,73] IC 95% - femmes Sp=0,73 (0,58-0,84] IC 95% - enfants Sp=0,76 (0,55-0,89] IC 95% 	<p>Hétérogénéité des études analysées en termes :</p> <ul style="list-style-type: none"> - prévalence d'appendicite - sex-ratio - variabilité (examinateur dépendant) des résultats des examens cliniques <p>Absence de données de suivi chez les patients non opérés, source de biais qui peut augmenter artificiellement r les valeurs de sensibilité et de spécificité.</p>

			<p>rétrospectives 1 étude design pas clair</p> <p>Articles exclus, portant exclusivement sur les femmes enceintes</p>		<p>Analyse du score en termes de rapports de risque :</p> <p>Rapport de risque : RR= nombre prédit par le score/nombre réel</p> <p>Hommes</p> <ul style="list-style-type: none"> - Score faible (1 à 4) : RR=1,06 (0,87-1,28] IC 95% - score moyen (5 à 6) : RR= 1,09 (0,86-1,37] IC 95% - score élevé (7 à10) : RR=1,02 (0,97-1,08] IC 95% <p>femmes</p> <ul style="list-style-type: none"> - Score faible (1 à 4) : RR= 5,35 (2,17-13,9] IC 95% - score moyen (5 à 6] : RR= 1,82 (1,20-2,78] IC 95% - score élevé (7 à10) : RR= 1,14 (1,04-1,25] IC 95% <p>Conclusion</p> <p>Le score d'Alvarado est un bon outil de tri pour orienter la décision vers le retour à domicile, l'observation ou la réalisation d'examens d'imagerie avant chirurgie</p>	
--	--	--	---	--	---	--

Méta-analyse de Ohle et al., 2011 (57)

Évaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori est fourni.
- La sélection des études de même que l'extraction des données est réalisée par deux personnes
- La recherche documentaire est réalisée sur 5 data bases.
- Il n'est pas précisé si le type de publication constitue un critère d'inclusion.
- La liste des études incluses n'est pas fournie.
- Les caractéristiques des études incluses sont indiquées.
- La qualité scientifique des études incluses est évaluée.
- La qualité scientifique des études incluses dans la revue est utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites et appropriées
- La probabilité d'un biais de publication n'a pas été évaluée.
- L'absence de conflits d'intérêts a été déclarée.

Tableau 20. Score d'Andersson, étude prospective Andersson & Andersson, 2008, (24)

Auteur, année, référence, Type d'étude	Objectif de l'étude	Population* Caractéristiques des patients	Critères de jugement	Données recueillies	Résultats et conclusions	Limites méthodologiques
Andersson & Andersson, 2008, (24) prospective consécutives	Construire et valider un système de score clinico-biologique en intégrant les marqueurs inflammatoires (CRP), sur des données recueillies de manière prospective sur 3 séries rétrospectives de patients avec suspicion d'appendicite.	Patients présentant des signes et symptômes cliniques faisant suspecter une appendicite. 545 patients Age moyen 25,9 ans	Résultat histopathologique après chirurgie suivi 1 mois patients non opérés	- Signes et symptômes cliniques - Résultats biologiques	Prévalence appendicite 36% Mesure du nombre total de leucocytes chez 274 patients. Leucocytes $>10. 10^9/L$: Se 76 % (65%-84 %) IC 95%, Sp 52% (45%-60%) IC 95% RVP= 1,59 (1,31-1,93] IC 95% RVN=0,46 (0,31-0,67] IC 95% Courbes ROC, AUC=0,72 (0,65-0,79] IC 95%. Le seuil optimal, à $11,9 \times 10^9$ cellules/L avec Se=69% et Sp=72% Température mesurée chez 293 patients $>99^\circ F$ soit $37,22^\circ C$ Se=47% (36%-57%) IC 95% Sp=64% (57%-71%) IC 95% RVP= 1,3 (0,97-1,72] IC 95% RVN= 0,82 (0,65-1,01] IC 95% AUC ROC=0,59 (0,52-0,66] IC 95%. Considérés chacun individuellement, le nombre total de leucocytes de même que la température corporelle ne permettent une discrimination entre les patients avec et sans appendicite	Données monocentriques Recueillies rétrospectivement sur une série prospective de patients consécutifs

Tableau 21. Comparaison de la performance diagnostique du scanner et de l'échographie méta-analyse de Van Randen *et al* (58)

Auteur, année, référence, pays	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Critères de jugement	Résultats Et conclusion	Limites méthodologiques
van Randen <i>et al.</i> , 2008 (58) Méta-analyse	1-Comparer la performance diagnostique du scanner et de l'échographie à compression progressive chez le même patient 2-Calculer les probabilités post-test obtenues avec le scanner et l'échographie sur 3 sous-groupes de patients avec des prévalences différentes : - Douleur abdominal aiguë (DDA) - Douleur en fosse iliaque droite (FID) Suspicion d'appendicite	Période: janvier 1966-février 2006 Bases : Medline, EMBASE, CINAHL, Cochrane. Mots clés renseignés : douleur abdominale, échographie e scanner.	Critères de sélection - Etude prospective comparant sur chaque patient de l'étude les deux modalités d'imagerie : scanner et l'échographie à compression progressive - Etudes de plus de 10 patients - Résultats présentés de manière à pouvoir constituer un tableau à 2 entrées pour le calcul des paramètres diagnostiques Résultats de la sélection - 6 études prospectives sont retenues pour le	- Diagnostic histopathologique pour les patients opérés - Suivi clinique pour les patients non opérés	<u>Caractéristiques des patients</u> : - Age moyen : de 26 à 38 ans avec une majorité autour de 38 ans (3 études) - Sexe-ratio F/H : de 0,32 et 2,84 - Présentation clinique de l'appendicite : 4 études patients avec suspicion d'appendicite (540 patients /671) et 2 études patients atypiques (131 patients) - Prévalence : entre 13 et 77% selon les études - Durée des suivi patients non opérés : de 1 à 13 mois selon les études <u>Sensibilité</u> Scanner 0,91(0,84-0,95] vs échographie 0,78 (0,67-0,86] p<0,017 <u>Spécificité</u> Scanner 0,90 (0,85-0,94] vs échographie 0,83 (0,76-0,88] p<0,037 <u>RVP</u> Scanner 9,29 (6,86-12,6] vs échographie 4,5 (3,03-6,68] p=0,011 <u>RVN</u> Scanner 0,10 (0,06-0,17] vs échographie 0,27(0,17-0,43] p=0,013 Taux de faux positifs dans les études : de 2 à 11 % avec le scanner et de 1 à 22% avec l'échographie. <u>Probabilités post test</u> <i>Patients >65 ans avec DAA</i> - Probabilité pré-test : 3% - Probabilité post test si test positif : scanner 22%, échographie 12% - Probabilité post-test si test négatif : scanner 0%,	hétérogénéité entre les études sur : - sexe-ratio - prévalence - la durée de suivi des patients non opérés Pas de test d'hétérogénéité décrit Biais de vérification Biais d'absence d'insu

			<p>calcul des paramètres diagnostiques Se, Sp, RVP, RVN Total 671 patients</p> <p>- 17 études sont éligibles pour le calcul de probabilité post-test en fonction de la prévalence Total 5093 patients</p>	<p>échographie 1%</p> <p><i>Patients avec DAA</i></p> <ul style="list-style-type: none"> - Probabilité pré-test : 6% - Probabilité post test si test positif : scanner 37%, échographie 22% - Probabilité post-test si test négatif : scanner 1%, échographie 2% <p><i>Patients avec suspicion d'appendicite ou douleur en FID</i></p> <ul style="list-style-type: none"> - Probabilité pré-test : 37% - Probabilité post test si test positif : scanner 85%, échographie 73% - Probabilité post-test si test négatif : scanner 6%, échographie 14% - <i>Toute la population de la méta-analyse</i> - Probabilité pré-test : 50% - Probabilité post test si test positif : scanner 90%, échographie 82% - Probabilité post-test si test négatif : scanner 9%, échographie 21% <p><u>Conclusion des auteurs</u></p> <p>Quelle que soit la probabilité pré-test, la comparaison de la performance diagnostique des 2 modalités d'imagerie chez les mêmes patients montre que le scanner est meilleur que l'échographie à compression progressive.</p> <p>Le taux de faux positifs reste important avec les deux modalités, il est moins important avec le scanner. Mais quelle que soit la modalité ce taux varie beaucoup selon les études, les auteurs attribuent cette différence à la variabilité inter-observateur.</p> <p>Même si le scanner est plus performant, les auteurs recommandent l'échographie en première intention chez la femme jeune et chez les personnes minces.</p>	
--	--	--	---	--	--

Méta-analyse de van Randen et al., 2008 (58)

Évaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori est fourni.
- La sélection des études de même que l'extraction des données a été faite par deux personnes
- La recherche documentaire est exhaustive réalisée sur 4 data-bases.
- Le type de publication constitue un critère d'inclusion.
- La liste des études incluses est fournie.
- Les caractéristiques des études incluses sont indiquées partiellement.
- La qualité scientifique des études incluses a été évaluée.
- La qualité scientifique des études incluses dans la revue n'est pas utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites et sont appropriées.
- La probabilité d'un biais de publication n'a pas été évaluée.
- La présence ou l'absence de conflits d'intérêts n'ont pas été déclarées.

Tableau 22. Evaluation de la performance diagnostique du scanner et de l'échographie Méta-analyse de Doria *et al* (76)

Auteur, année, référence, pays	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Critères de jugement	Résultats Et conclusion	Limites méthodologiques
Doria, 2006 (76)	Evaluer la performance diagnostique de l'échographie et de la tomodensitométrie (CT) pour le diagnostic de l'appendicite en pédiatrie et chez les populations adultes	<p>Période de recherche : janvier 1986- décembre 2004</p> <p>Bases : Medline, EMBASE, CINAHL, Cochrane controlled trial registered, the Cochrane database of systematic reviews, American College of Physicians Journal Club databases.</p> <p>Mots clés renseignés : appendicitis, appendix, computed, tomography scan, computed tomography sonography</p> <p>Pas de restriction sur la langue</p>	<p>Critères d'inclusion des études</p> <ul style="list-style-type: none"> - Etude prospective et rétrospectives évaluant la performance diagnostique du scanner et ou de l'échographie - Critères de jugement : résultat histopathologique et suivi clinique de plus d'une semaine ; - Disponibilité des données : vrais positifs, faux positifs, vrais négatifs et faux négatifs - Segmentation des résultats en fonction de l'âge avec un maximum de 20 ans pour la population pédiatrique et un minimum de 13 ans pour la population adulte. - Critères de positivité à l'imagerie définis par un diamètre >6mm, appendice non compressible à l'échographie - En cas de non visualisation de l'appendice, présence de signes inflammatoires - Description des techniques scanner et de la modalité de contraste - Ratio F/M <3/1 	<ul style="list-style-type: none"> - Diagnostic histopathologique pour les patients opérés - Suivi clinique pour les patients non opérés plus d'une semaine. 	<p><u>Caractéristiques des patients</u> :</p> <ul style="list-style-type: none"> - Sex-ratio F/H : non renseigné - Moyenne d'âge de 20 à 49 ans - Prévalence moyenne 48% avec l'échographie et 40% avec le scanner ; - Le taux de perforation 18,5 % (dans 3 études) <p><u>Résultats échographie</u></p> <p>13 études (12 prospectives et 1 rétrospective)</p> <p>1579 patients</p> <p>Sensibilité= 83% (78-87] IC 95%</p> <p>Spécificité= 94% (92-95] IC 95%</p> <p>RVP*= 13,8</p> <p>RVN*=0,18</p> <p><u>Résultats scanner</u></p> <p>18 études (13 prospectives et 5 rétrospective)</p> <p>2439 patients</p> <p>Sensibilité= 94% (92-95] IC 95%</p> <p>Spécificité= 94% (94-96%] IC 95%</p> <p>RVP*= 15,66</p> <p>RVN*=0,06</p> <p><u>Conclusion des auteurs</u></p> <p>Sur la base des résultats de cette méta-analyse, le scanner semble avoir une sensibilité significativement plus élevée que l'échographie pour le diagnostic d'appendicite aiguë.</p>	<p>Limites inhérentes à la qualité des études analysées :</p> <ul style="list-style-type: none"> - Biais de vérification - Absence d'insu - Absence de description de la présentation cliniques et de la gravité des signes - Absence de données sur le poids et sur l'IMC - Absence de randomisation pour la comparaison de l'échographie et du scanner, - Difficulté de définir un seuil de positivité pour les examens équivoques

			<p>- Etudes avec une prévalence d'appendicite comprise entre 15 et 75%.</p> <p><i>Critères d'exclusion</i></p> <p>- Etudes de cas, série de cas, revues, études avec données non publiées, abstracts et lettres à l'éditeur</p> <p>- Objectifs de l'étude autre que l'évaluation de la performance diagnostique</p> <p>- Les études doublons</p>			
--	--	--	--	--	--	--

* calculé à partir des valeurs de Se et Sp

Méta-analyse de Doria, 2009 (76)

Evaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori est fourni.
- La sélection des études de même que l'extraction des données a été faite par deux personnes
- La recherche documentaire est exhaustive.
- Seules les études prospectives et rétrospectives ont été sélectionnées.
- La liste des études incluses est fournie sur un document annexe.
- Les caractéristiques des études incluses sont partiellement indiquées.
- La qualité scientifique des études incluses a été évaluée par deux lecteurs indépendants.
- La qualité scientifique des études incluses dans la revue n'est pas utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites et semblent appropriées.
- La probabilité d'un biais de publication est évaluée.
- La présence ou l'absence de conflits d'intérêts n'ont pas été déclarées.

Tableau 23. Evaluation de la performance diagnostique du scanner sans contraste revue systématique Hlibczuk *et al.* 2010, (60)

Auteur, année, référence, pays	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Critères de jugement	Résultats Et conclusion	Limites Niveau de preuves
Hlibczuk <i>et al.</i> , 2010 (60)	L'objectif de cette revue systématique est d'évaluer la preuve pour l'utilisation du scanner sans pour le diagnostic d'appendicite aiguë chez les adultes	<p>Période: janvier 1950-mars 2008</p> <p>Bases : Medline, EMBASE, Cochrane library, Health technology database, Mise à jour jusqu'en mars 2009</p> <p>Mots clés renseignés : appendicitis, appendectomy, x-ray computed, contrast media, computed tomography, helical, enhanced, contrast, non contrast..</p>	<p>Critères d'inclusion</p> <ul style="list-style-type: none"> - patients admis aux urgences pour douleur abdominale aiguë avec suspicion d'appendicite - patients âge \geq 16 ans - si la référence standard pour le diagnostic ou l'exclusion de l'appendicite aiguë est utilisé chez tous les patients ; - les études sont retenues si elle répondent à 6 (6/11) items de la grille QUADAS, définis comme les plus pertinents par les auteurs <p>Critères d'exclusion</p> <ul style="list-style-type: none"> - Etudes qui compilent des populations adulte et pédiatrique - études visant à évaluer la précision des scanners monobarettes - Etudes avec un biais de vérification partielle - si la période du suivi clinique pour les patients non opérés n'est pas indiquée ou n'a pas pu être confirmée par contact de l'auteur <p>Résultats de la sélection</p> <ul style="list-style-type: none"> - 7 études prospectives Total 1060 patients 	<ul style="list-style-type: none"> - Diagnostic macroscopique par le chirurgien - Diagnostic histopathologique pour les patients opérés - Suivi clinique pour les patients non opérés minimum 2 semaines. 	<p><u>Caractéristiques des patients</u> :</p> <ul style="list-style-type: none"> - Fourchette d'âge comprise entre 16 et 86 ans avec 4 études avec un âge maximal autour de 65 ans et 3 études avec un âge max autour de 82 ans. - Sex-ratio F/H : non renseigné - Prévalence médiane 39,3% (entre 20,1 et 84,5% selon les études) <p><u>Résultats combinés</u></p> <p>Sensibilité= 92,7% (89,5-95] IC 95%</p> <p>Spécificité= 96,1% (94,2-97] IC 95%</p> <p>RVP= 24</p> <p>RVN=0,08</p> <p>Au regard des résultats compilés de spécificité, si l'imagerie scanner est généralisée, le taux acceptable d'appendicite négatives devrait être de 4%.</p> <p><u>Conclusion</u></p> <p>Pour le diagnostic d'appendicite aiguë chez une population adulte, la précision diagnostique du scanner sans produit de contraste est suffisante pour permettre la prise de décision clinique au sein d'un service d'urgence</p>	<ul style="list-style-type: none"> - Biais de publication non vérifié - hétérogénéité de prévalence d'appendicite entre les études, particulièrement élevée dans 2 études (77,6 et 84,5 %] - Biais de vérification dans 6 études sur 7 de la méta-analyse. - Le taux de résultats d'examen de scanner non interprétable n'est rapporté que dans 2 études. Les autres études considèrent ces examens comme négatifs.

Revue systématique de Hlibczuk et al., 2010, (60)

Evaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori est fourni.
- La sélection des études de même que l'extraction des données a été faite par deux personnes
- La recherche documentaire est exhaustive 3 data-bases.
- Le type de publication ne constituait pas un critère d'inclusion..
- La liste des études incluses est fournie est fournie.
- Les caractéristiques des études incluses sont indiquées.
- La qualité scientifique des études incluses a été évaluée par deux lecteurs indépendants à l'aide de la grille QUADAS.
- La qualité scientifique des études incluses dans la revue n'est pas utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites et semblent appropriées.
- La probabilité d'un biais de publication n'a pas été évaluée.
- La présence ou l'absence de conflits d'intérêts n'ont pas été déclarées.

Tableau 24. Evaluation de la performance diagnostique de l'échographie et du scanner dans la détection de l'appendicite aiguë Terasawa *et al.* 2004, (48)

Auteur, année, référence, pays	Objectif de l'étude	Stratégie de recherche bibliographique	Critères de sélection des articles	Critères de jugement	Résultats Et conclusion	Limites Niveau de preuves
<p>Terasawa <i>et al.</i>, 2004 (48)</p>	<p>Définir la précision diagnostique de la tomodensitométrie et de l'échographie pour le diagnostic de l'appendicite chez les adultes et les adolescents, en répondant aux 2 questions suivantes :</p> <p>1) Quelle est la précision diagnostique de la tomodensitométrie et de l'échographie ?</p> <p>2) Quelles sont les forces et les limites de la littérature actuelle</p>	<p>Période : 1966-décembre 2003</p> <p>Bases : Medline, EMBASE,</p> <p>Termes Mesh utilisés : renseignements : appendicitis, appendix, x-ray computed, tomography, ultrasonography, sensitivity, diagnostic use, diagnosis</p>	<p>Critères d'inclusion</p> <ul style="list-style-type: none"> - Etudes prospectives qui évaluent le scanner ou l'échographie à compression progressive chez les adultes et les adolescents avec suspicion d'appendicite - Confirmation par l'examen histopathologique après chirurgie ou par suivi clinique chez les patients non opérés. - patients âge ≥ 14 ans <p>Résultats de la sélection</p> <ul style="list-style-type: none"> - 9 séries prospectives pour le scanner - 13 séries prospectives pour l'échographie - 1 étude contrôlée randomisée comparant l'échographie au scanner - 2 études contrôlées 	<ul style="list-style-type: none"> - Diagnostic histopathologique pour les patients opérés - Suivi clinique pour les patients non opérés <p>Le seuil de positivité adopté aussi bien pour les études scanner que échographie est un diamètre de l'appendice supérieur à 6 mm</p>	<p><u>Caractéristiques des études :</u></p> <p><i>Etudes scanner</i></p> <ul style="list-style-type: none"> - N=1172 patients dont 504 (43%) avec suspicion d'appendicite et 668 (57%) de présentations atypiques - Age moyen comprise entre 31 et 38 ans - Pourcentage de femmes 61% - Sex-ratio F/H calculé : 1,56 - Prévalence moyenne 45% (de 24 et 78 % selon les études) - Taux moyen d'appendicites négatives 36% (de 14 à 47% selon les études) - Diversité es techniques scanner : helicoidal, non helicoidal, abdominopelvien ou focalisé avec contraste (oral, IV, rectal, oral+IV, ou plusieurs modalités à la fois) <p><i>Etudes échographie</i></p> <ul style="list-style-type: none"> - N=1516 patients dont 1296 (85%) avec suspicion d'appendicite et 220 (14,5%) de présentations atypiques - Age moyen comprise entre 20 et 46 ans - Pourcentage de femmes 55% - Sex-ratio F/H calculé : 1,22 - Prévalence moyenne 50% (de 13 et 75 % selon les études) - Taux moyen d'appendicites négatives 36% (calculé uniquement sur 10 études varie de 14 à 50%) 	<p>Faible qualité méthodologique des études originales :</p> <ul style="list-style-type: none"> - Biais de vérification différentiel - Biais d'absence d'insu pour l'imagerie - Biais d'absence d'insu pour la réalisation du test de référence - Absence de description du degré de gravité des signes : biais de spectre - Hétérogénéité importante entre les études notamment pour l'échographie - Grande variabilité de la prévalence d'appendicite entre les études

			<p>randomisées comparant le scanner à l'examen clinique</p>	<p>– Diversité es techniques d'échographie : sondes linéaires ou courbes ou linéaires courbes. plusieurs sondes avec des fréquences différentes : de 2,5 à 9,0 MHz</p> <p><u>Résultats de la méta-analyse</u></p> <p><i>Scanner</i></p> <p>Sensibilité= 0,94 (0,9-0,95] IC 95%</p> <p>Spécificité= 0,95 (0,93-0,96] IC 95%</p> <p>RVP= 13,3 (9,9-17,9]</p> <p>RVN=0,09 (0,07-0,13]</p> <p><i>Echographie</i></p> <p>Sensibilité= 0,86 (0,83-0,88] IC 95%</p> <p>Spécificité= 0,81 (0,78-0,84] IC 95%</p> <p>RVP= 5,8 (3,5-9,5]</p> <p>RVN=0,19 (0,13-0,27]</p> <p><u>Résultats de la méta-analyse des études comparatives scanner vs échographie</u></p> <p>Résultats obtenus sur 4 études</p> <p>RVP= 12,2 (7,1-21,2] pour le scanner vs 3,4 (2,4-4,8] pour l'échographie</p> <p>RVN=0,05 (0,02-0,12] pour le scanner vs 0,33 (0,19-0,55] pour l'échographie).</p> <p><u>Conclusion</u></p> <p>Les résultats de la méta-analyse démontrent que le scanner a une très bonne performance diagnostique qui par ailleurs semble être supérieure à celle de l'échographie particulièrement chez l'adulte et l'adolescent. Néanmoins les faiblesses méthodologiques des études originales pourraient induire une surestimation de la précision diagnostique aussi bien pour le scanner que pour l'échographie.</p>
--	--	--	---	---

Méta-analyse de Terasawa et al., 2004, (48)

Évaluation de la qualité de cette étude selon la grille AMSTAR (outil d'évaluation de la qualité des revues systématiques) :

- Le plan de recherche établi a priori est fourni.
- La sélection des études de même que l'extraction des données a été faite par deux personnes
- La recherche documentaire est exhaustive 2 data-bases.
- Seules les études prospectives sont incluses
- La liste des études incluses est fournie est fournie.
- Les caractéristiques des études incluses sont indiquées.
- La qualité scientifique des études incluses a été évaluée par deux lecteurs indépendants (méthode Cochrane)
- La qualité scientifique des études incluses dans la revue est utilisée adéquatement dans la formulation des conclusions.
- Les méthodes utilisées pour combiner les résultats des études sont décrites et semblent appropriées.
- La probabilité d'un biais de publication n'a pas été évaluée.
- L'absence de conflits d'intérêts est déclarée.

Annexe 10. Analyse des réponses après cotation

Proposition soumise au vote	Nombres de réponses par valeur de 1 (totalement inappropriée) à 9 (totalement appropriée)									Valeur médiane	Nombre de cotateurs	Nombre de valeurs manquantes	Jugement retenu
	1	2	3	4	5	6	7	8	9				
Proposition 1					1			4	12	9	17	0	Appropriée avec un accord fort
Proposition 2							1	4	12	9	17	0	Appropriée avec un accord fort
Proposition 3			1			1		5	10	9	17	0	Appropriée avec un accord relatif
Proposition 4							5	2	10	9	17	0	Appropriée avec un accord fort
Proposition 5			2			2		2	9	9	17	0	Appropriée avec un accord relatif
Proposition 6							1	6	9	9	16*	0	Appropriée avec un accord fort
Proposition 7						1	1	4	10	9	16*	0	Appropriée avec un accord relatif
Proposition 8		1		1		2	3	1	7	8	15*	1	Affirmation incertaine
Proposition 9			1			3	4	2	5	7	15*	1	Appropriée avec un accord relatif
Proposition 10					2	1	1	3	9	9	16	1	Appropriée avec un accord relatif

*Un expert biologiste avait informé « *a priori* » qu'il ne répondrait qu'aux questions relevant de ses compétences. Il n'a pas répondu aux questions concernant les examens morphologiques, par conséquent sa voix n'a pas été prise en compte pour ces questions.

Annexe 11. Recueil de l'avis des professionnels sur la qualité de l'argumentaire scientifique

Concernant l'argumentaire scientifique, les experts du groupe d'orientation ont été sollicités pour apprécier la lisibilité, la pertinence des informations analysées, la rigueur méthodologique et l'adéquation des conclusions avec les données analysées.

A la question « *Souhaitez-vous émettre un commentaire général concernant le rapport provisoire d'évaluation que vous avez reçu ? (lisibilité, pertinences des informations analysées, rigueur méthodologique, adéquation des conclusions avec les données analysées...)* » 7 experts sur 17 ont émis les commentaires suivants :

- travail d'une grande qualité, très lisible, clair, avec une analyse bibliographique sérieuse, dont les résultats sont en accord avec la pratique quotidienne d'un centre d'urgence. L'arbre décisionnel paraît bien résumer l'ensemble des données fournies, et peut être appliqué dans tout centre disposant d'imagerie ;
- bonne lisibilité du rapport, quelques fautes à corriger ;
- ce rapport m'apparaît très fondé ;
- le travail bibliographique et l'analyse qui en est faite sont très bien faits et présentés rendant la lecture du rapport aisé. Basé uniquement sur des travaux dont le niveau de preuve est élevé, ce document établit une vision assez claire de ce qui actuellement prouve ce qui justifie mon accord total avec les 8 premières propositions du questionnaire (je ne suis pas concerné par les propositions 9 et 10). La contrepartie de ce travail statistique est d'aboutir à un algorithme très théorique un peu en décalage avec la réalité quotidienne : combien de patients sans défense mais avec une vraie sensibilité de la FID ou une douleur migrante sortiront sans imagerie ? Combien de patients avec une triade seront opérés sans avoir éliminé un abcès, une maladie de Crohn ?
- la critique est aisée mais l'art est difficile, vous critiquez systématiquement toutes les études, méta-analyse ou Cochrane pourtant faite par des méthodologistes certes non français mais néanmoins de haut niveau. Je ne trouve pas qu'il y ait si peu de corpus de littérature ; il y a beaucoup d'études et d'essais randomisés sur votre sujet avec méta-analyse et recommandations, comparé à la pathologie infectieuse gynécologique par exemple recommandations qui sont en cours et que je viens d'avoir à commenter !? Il faut donc faire avec ce que l'on a, vos conclusions me semblent toujours négatives alors que d'autres ont statué même si ce n'est pas parfait et c'est ce que l'on nous demande. Il me paraît aussi logique que les experts de l'HAS avec le groupe d'orientation et de celui de cotation proposent les études pour améliorer les niveaux de preuve et participent à leur mise en œuvre et à leur demande de financement par le biais de la sélection des PHRC ou autres : exemple pour les scores à valider, le TDM et la coelioscopie avec les différents sous groupes, selon le sexe l'âge et l'obésité ;
- n'est ce pas aussi l'occasion de faire le point de nos connaissances et d'orienter les recherches en précisant les points les plus incertains. Par exemple la place des antibiotiques ne me paraît pas du tout claire mais potentiellement importante ;
- il aurait été intéressant d'inclure dans la bibliographie des études analysant l'intérêt de l'échographie endovaginale couplée ou non à l'échographie transpéritéale en cas de douleur abdominale chez la femme jeune et la comparant avec les performances de l'échographie transpéritéale ? En effet, celle-ci n'a pas d'intérêt pour le diagnostic positif d'appendicite aiguë mais elle a une valeur prédictive négative importante en raison de sa performance dans les diagnostics différentiels de la douleur abdominale chez la femme jeune.

Annexe 12. Liste des tableaux

Tableau 1. Signes et symptômes cliniques, capacité discriminante (AUC) et performance diagnostique (RVP et RVN).....	29
Tableau 2. Caractéristiques des patients dans les études.....	31
Tableau 3. Nombre de leucocytes, capacité discriminante (AUC) et performance diagnostique (RVP et RVN).....	34
Tableau 4. Nombre de granulocytes, capacité discriminante (AUC) et performance diagnostique (RVP et RVN).....	34
Tableau 5. Taux PNN, capacité discriminante (AUC) et performance diagnostique (RVP et RVN).....	35
Tableau 6. CRP, capacité discriminante (AUC) et performance diagnostique (RVP et RVN).....	35
Tableau 7. Variables combinées, capacité discriminante (AUC) et performance diagnostique (RVP et RVN).....	37
Tableau 8. Les différents profils clinico-biologiques permettant d'obtenir une prévalence d'appendicites \geq 75% (47).....	39
Tableau 9. Valeurs agrégées de sensibilité et de spécificité aux valeurs de score 5 et 7(57).....	41
Tableau 10. Rapports de risque en fonction du niveau de score et en fonction de l'âge et du sexe (57).....	41
Tableau 11. Valeur diagnostique de l'échographie pour le diagnostic de l'appendicite chez l'adulte, résultats de méta-analyses.....	49
Tableau 12. Valeur diagnostique du scanner pour le diagnostic de l'appendicite chez l'adulte, résultats de méta-analyses.....	50
Tableau 13. Comparaison directe des résultats de performance diagnostique du scanner et de l'échographie.....	55
Tableau 14. Méta-analyses sur les symptômes et signes cliniques R.E.B Anderson, 2004 (51).....	108
Tableau 15. Revue systématique sur les symptômes et signes cliniques James M. Wagner 1996, (52).....	111
Tableau 16. Etude prospective de Laméris <i>et al.</i> (47).....	113
Tableau 17. Etude prospective de Cardall <i>et al</i> (54).....	115
Tableau 18. Etude prospective d'Ortega-Deballon <i>et al</i> (46).....	116
Tableau 19. Méta-analyse sur le score d'Alvaradon Robert Ohle <i>et al</i> 2011 (57).....	118
Tableau 20. Score d'Andersson, étude prospective Andersson & Andersson, 2008, (24).....	121
Tableau 21. Comparaison de la performance diagnostique du scanner et de l'échographie méta-analyse de Van Randen <i>et al</i> (58).....	122
Tableau 22. Evaluation de la performance diagnostique du scanner et de l'échographie Méta-analyse de Doria <i>et al</i> (76).....	125
Tableau 23. Evaluation de la performance diagnostique du scanner sans contraste revue systématique Hlibczuk <i>et al.</i> 2010, (60).....	127
Tableau 24. Evaluation de la performance diagnostique de l'échographie et du scanner dans la détection de l'appendicite aiguë Terasawa <i>et al.</i> 2004, (48).....	129

Participants

► Groupe d'orientation

Dr Hervé BERCHE	Généraliste, Compiègne (60)
Dr Stéphanie FRANCHI-ABELLA	Radiopédiatre, CHU du Kremlin Bicêtre (94)
Dr Philippe LEVEAU	Urgentiste, Centre Hospitalier Nord 2 Sèvres, Thouars (79)
Pr Philippe MARTEAU	Hépatogastro-entérologue, Hôpital Lariboisière, Paris (75)
Pr Christophe PONCELET	Gynécologue-Obstétricien, CHU Jean-Verdier, Bondy (93)
Dr Nathalie SIAUVE	Radiologue, Hôpital Européen Georges Pompidou, Paris (75)
Pr Karem SLIM	Chirurgien digestif, CHU Estaing, Clermont-Ferrand (63)
Pr Benoit TERRIS	Anatomo-pathologiste, Hôpital Cochin, Paris (75)
Pr Bertrand TOUSSAINT	Médecin biologiste, CHU de Grenoble, (38)

► Groupe de cotation

Pr Claude BENDAVID	Biologiste, CHU de Rennes (35)
Pr Jacques BIENVENU	Biologiste, Université Claude Bernard, les hospices civils de Lyon (69)
Dr Marie-Jeanne BOUDET	Chirurgien digestif, Clinique Labrouste, Paris (75)
Dr André DABROWSKI	Chirurgien digestif, Clinique chirurgicale de Saint-Omer (62)
Pr Pierre DARNAULT	Radiologue, CHU de Rennes, Rennes (35)
Dr François ESCAT	Urgentiste, Clinique Occitanie, Muret (31)
Dr Eric FRAMPAS	Radiologue, CHU de Nantes, Hôtel Dieu, Nantes (44)
Dr Agnès LEURET	Urgentiste, Centre Hospitalier William Morey, Chalon sur Saône (71)
Dr Jean-Marc LORPHELIN	Radiologue, Hôpital privé de Bois Bernard, Bois Bernard (62)
Pr Henri MARRET	Gynécologue-obstétricien, Hôpital Bretonneau, Tours (37)
Dr Muriel MATHONNET	Chirurgien digestif, CHU de Limoges, Limoges (87)
Dr Bertrand MORIN	Chirurgien digestif, Groupe Hospitalier du Havre, Le Havre (76)
Dr Jean-Marc PAULY	Généraliste, Rodemack (57)
Dr Jacques REMIZE	Urgentiste, Centre Hospitalier, Brive La Gaillarde (19)
Dr Janick SELVES	Anatomo-pathologiste, Université Paul Sabatier, Toulouse (31)
Dr Jean SENDE	Urgentiste, Hôpital Privé Armand Brillard, Nogent sur Marne (94)
Dr Magaly ZAPPA	Radiologue, Hôpital Beaujon, Clichy (92)

Références

1. Vigneron E. Pertinence des actes. Césariennes et appendicectomies. *Rev Hosp Fr* 2008;(525):59-62.
2. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986;2(1):53-63.
3. What do we mean by appropriate health care? Report of a working group prepared for the Director of Research and Development of the NHS Management Executive. *Qual Health Care* 1993;2(2):117-23.
4. Haute Autorité de Santé. Appendicectomie. Eléments décisionnels pour une indication pertinente. Note de cadrage. Saint-Denis La Plaine: HAS; 2011.
http://www.has-sante.fr/portail/plugins/ModuleXitiKLEE/types/FileDocument/oXiti.jsp?id=c_1218030
5. Silen W. Appendicite aiguë. In: Fauci AS, Braunwald E, Isselbacher KJ, Wilson JD, Martin JB, Kasper DL, et al., ed. *Harrison. Médecine interne*. Maidenhead: McGraw-Hill; 2000. p. 1911-1914.
6. Latulippe LG, Demers P. Le diagnostic de l'appendicite aiguë par l'omnipraticien : le défi est de taille! *Clinicien* 2002;17(10):107-15.
7. Société nationale française de gastroentérologie. Item 224. Appendicite de l'enfant et de l'adulte 2009. <http://www.snfge.org/05-Interne-Chercheurs/0B-internes-etudiants/masson/chapitre_224.pdf> .
8. Flamant Y, Associations de recherche en chirurgie. Douleurs abdominales aiguës de l'adulte. In: Rambaud JC, ed. *Traité de gastro-entérologie*. Paris: Flammarion; 2005. p. 198-210.
9. Mathias J, Bruot O, Ganne PA, Laurent V, Regent D. Appendicite. *Encycl Méd Chir Radiodiagnostic - Appareil digestif* 2008;33-472-G-10.
10. Société nationale française de gastroentérologie. Appendicite. Reims: SNFGE; 2001.
<http://www.snfge.org/01-Bibliotheque/0D-Pratiques-cliniques/RPC-appendic.pdf>
11. Rohr S, Lang H, Mechine A, Meyer C. Appendicite aiguë. *Encycl Méd Chir Gastro-entérologie* 1999;9-066-A-10.
12. Agence nationale pour le développement de l'évaluation médicale. Evaluation des méthodes coelioscopiques en chirurgie digestive. Paris: ANDEM; 1994.
<http://www.has-sante.fr/portail/upload/docs/application/pdf/coeliodigestive%20.pdf>
13. Craig S. *Appendicitis, acute*. New York: Medscape; 2010.
14. Varadhan KK, Humes DJ, Neal KR, Lobo DN. Antibiotic therapy versus appendectomy for acute appendicitis: a meta-analysis. *World J Surg* 2010;34(2):199-209.
15. Simillis C, Symeonides P, Shorthouse AJ, Tekkis PP. A meta-analysis comparing conservative treatment versus acute appendectomy for complicated appendicitis (abscess or phlegmon). *Surgery* 2010;147(6):818-29.
16. Vons C, Barry C, Maitre S, Pautrat K, Leconte M, Costaglioli B, et al. Amoxicillin plus clavulanic acid versus appendicectomy for treatment of acute uncomplicated appendicitis: an open-label, non-inferiority, randomised controlled trial. *Lancet* 2011;377(9777):1573-9.
17. Mason RJ. Appendicitis: is surgery the best option? *Lancet* 2011;377(9777):1545-6.
18. Mutter D, Marescaux J. Appendicites aiguës. Item 224. Module 14 2000. <http://udsmmed.u-strasbg.fr/emed/courses/MODULE14/document/Appendicites_aigues.pdf?cidReq=MODULE14> .
19. Société nationale française de gastroentérologie. Urgence : Appendicite 1999. <http://www.snfge.asso.fr/02-Connaitre-maladie/OH-urgences/faq/urgences_appendicite.htm> .
20. Wind P, Malamut G, Cuénod C-A, Bénichou J. Stratégie des explorations des douleurs abdominales. *Encycl Méd Chir Médecine d'urgence* 2007;25-050-A-20.
21. Menegaux F. Appendicites aiguës. In: Carli P, Riou B, Télion C, ed. *Urgences médico-chirurgicales de l'adulte*. Paris: Arnette Blackwell; 2011. p. 486-490.
22. Mondor H. *Diagnostics urgents. Abdomen*. Paris: Masson; 1947.
23. Montali I, von Flüe M. L'appendicite aiguë aujourd'hui. Nouveaux aspects d'une maladie connue de longue date. *Forum Med Suisse* 2008;8(24):451-5.
24. Andersson M, Andersson RE. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. *World J Surg* 2008;32(8):1843-9.
25. Brigand C, Steinmetz J-P, Rohr S. De l'intérêt des scores en matière de diagnostic d'appendicite. *J Chir Visc* 2009;146(5S1):2-7.
26. Haute Autorité de Santé. Principales indications et "non-indications" de la radiographie de l'abdomen sans préparation. Saint-Denis La Plaine: HAS; 2009.
http://www.has-sante.fr/portail/upload/docs/application/pdf/2009-02/rapport_rx_asp.pdf
27. Panebianco NL, Jahnes K, Mills AM. Imaging and laboratory testing in acute abdominal pain. *Emerg Med Clin North Am* 2011;29(2):175-93, vii.
28. Bretagnol F, Zappa M, Panis Y. Place de l'imagerie dans le diagnostic d'appendicite aiguë. *J Chir* 2009;146 Spec No 1:8-11.
29. Parks NA, Schroepel TJ. Update on imaging for acute appendicitis. *Surg Clin North Am* 2011;91(1):141-54.
30. Stoker J, van Randen A, Lameris W, Boermeester MA. Imaging patients with acute abdominal pain. *Radiology* 2009;253(1):31-46.

31. Société française de radiologie. Guide pratique à l'usage des médecins radiologues pour l'évaluation de leurs pratiques professionnelles. Paris: SFR; 2010.
32. Haute Autorité de Santé. Guide du bon usage des examens d'imagerie médicale. Saint-Denis La Plaine: HAS; 2005.
http://www.has-sante.fr/portail/upload/docs/application/pdf/examens_imagerie_medicale_guide_2006_12_28_17_11_41_41.pdf
33. Brenner DJ, Hall EJ. Computed tomography: an increasing source of radiation exposure. *N Engl J Med* 2007;357(22):2277-84.
34. Verdun FR, Bochud F, Gundinchet F, Aroua A, Schnyder P, Meuli R. Quality initiatives* radiation risk: what you should know to tell your patient. *Radiographics* 2008;28(7):1807-16.
35. Pedrosa I, Levine D, Eyvazzadeh AD, Siewert B, Ngo L, Rofsky NM. MR imaging evaluation of acute appendicitis in pregnancy. *Radiology* 2006;238(3):891-9.
36. American College of Radiology. Right lower quadrant pain. Suspected appendicitis. ACR appropriateness criteria. Reston: ACR; 2010.
<http://www.acr.org/-/media/7425A3E08975451EAB571A316DB4CA1B.pdf>
37. Flum DR, Morris A, Koepsell T, Dellinger EP. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. *JAMA* 2001;286(14):1748-53.
38. Agence nationale pour le développement de l'évaluation médicale. Appendicite. Recommandations et références médicales. *Gastroenterol Clin Biol* 1997;21(1):50-70.
39. Keyzer C, Cullus P, Tack D, de Maertelaer V, Bohy P, Gevenois PA. MDCT for suspected acute appendicitis in adults: impact of oral and IV contrast media at standard-dose and simulated low-dose techniques. *AJR Am J Roentgenol* 2009;193(5):1272-81.
40. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155(8):529-36.
41. Agence nationale d'accréditation et d'évaluation en santé. Guide d'analyse de la littérature et gradation des recommandations. Paris: ANAES; 2000.
<http://www.has-sante.fr/portail/upload/docs/application/pdf/analiterat.pdf>
42. Fédération Nationale des Centres de Lutte Contre le Cancer, Agence nationale d'accréditation et d'évaluation en santé, Fédération des médecins suisses. Grille d'évaluation de la qualité des recommandations pour la pratique clinique. Traduction française de la grille AGREE (Appraisal of Guidelines for Research and Evaluation Instrument). Paris: FNCLCC; 2002.
<http://www.has-sante.fr/portail/upload/docs/application/pdf/grille.pdf>
43. Agence nationale d'accréditation et d'évaluation en santé. Les conférences de consensus. Base méthodologique pour leur réalisation en France. Paris: ANAES; 1999.
<http://www.has-sante.fr/portail/upload/docs/application/pdf/guideCC.pdf>
44. Haute Autorité de Santé. Élaboration de recommandations de bonne pratique. Méthode « Recommandations par consensus formalisé ». Guide méthodologique. Saint-Denis La Plaine: HAS; 2010.
http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-01/guide_methodologique_consensus_formalise.pdf
45. Lintula H, Kokki H, Pulkkinen J, Kettunen R, Grohn O, Eskelinen M. Diagnostic score in acute appendicitis. Validation of a diagnostic score (Lintula score) for adults with suspected appendicitis. *Langenbecks Arch Surg* 2010;395(5):495-500.
46. Ortega-Deballon P, Ruiz de Adana-Belbel JC, Hernández-Matías A, García-Septiem J, Moreno-Azcoita M. Usefulness of laboratory data in the management of right iliac fossa pain in adults. *Dis Colon Rectum* 2008;51(7):1093-9.
47. Laméris W, van Randen A, Go PM, Bouma WH, Donkervoort SC, Bossuyt PM, et al. Single and combined diagnostic value of clinical features and laboratory tests in acute appendicitis. *Acad Emerg Med* 2009;16(9):835-42.
48. Terasawa T, Blackmore CC, Bent S, Kohlwes RJ. Systematic review: computed tomography and ultrasonography to detect acute appendicitis in adults and adolescents. *Ann Intern Med* 2004;141(7):537-46.
49. Pieper R, Kager L, Nasman P. Clinical significance of mucosal inflammation of the vermiform appendix. *Ann Surg* 1983;197(3):368-74.
50. Cobben LP, de Mol van Otterloo A, Puylaert JB. Spontaneously resolving appendicitis: frequency and natural history in 60 patients. *Radiology* 2000;215(2):349-52.
51. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. *Br J Surg* 2004;91(1):28-37.
52. Wagner JM, McKinney WP, Carpenter JL. Does this patient have appendicitis? *JAMA* 1996;276(19):1589-94.
53. American College of Emergency Physicians, Howell JM, Eddy OL, Lukens TW, Thiessen ME, Weingart SD, et al. Clinical policy: Critical issues in the evaluation and management of emergency department patients with suspected appendicitis. *Ann Emerg Med* 2010;55(1):71-116.
54. Cardall T, Glasser J, Guss DA. Clinical value of the total white blood cell count and temperature in the evaluation of patients with suspected appendicitis. *Acad Emerg Med* 2004;11(10):1021-7.
55. American College of Radiology. Acute abdominal pain and fever or suspected abdominal abscess. ACR appropriateness criteria®. ACR; 2008.
http://gm.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonGastrointestinalImaging/AcuteAbdominalPainandFeverorSuspectedAbdominalAbscessDoc1.aspx
56. Roth C, Tello R, Sutherland K, Ptak T. Prediction rule for etiology of vague abdominal pain in the emergency room: utility for imaging triage. *Invest Radiol* 2002;37(10):552-6.
57. Ohle R, O'Reilly F, O'Brien KK, Fahey T, Dimitrov BD. The Alvarado score for predicting acute appendicitis: A systematic review. *BMC Med* 2011;9(1):139.
58. van Randen A, Bipat S, Zwinderman AH, Ubbink DT, Stoker J, Boermeester MA. Acute appendicitis: meta-analysis

- of diagnostic performance of CT and graded compression US related to prevalence of disease. *Radiology* 2008;249(1):97-106.
59. Doria AS, Moineddin R, Kellenberger CJ, Epelman M, Beyene J, Schuh S, et al. US or CT for diagnosis of appendicitis in children and adults? A meta-analysis. *Radiology* 2006;241(1):83-94.
60. Hlibczuk V, Dattaro JA, Jin Z, Falzon L, Brown MD. Diagnostic accuracy of noncontrast computed tomography for appendicitis in adults: a systematic review. *Ann Emerg Med* 2010;55(1):51-9.
61. Horton MD, Counter SF, Florence MG, Hart MJ. A prospective trial of computed tomography and ultrasonography for diagnosing appendicitis in the atypical patient. *Am J Surg* 2000;179(5):379-81.
62. Hershko DD, Awad N, Fischer D, Mahajna A, Guralnik L, Israelit SH, et al. Focused helical CT using rectal contrast material only as the preferred technique for the diagnosis of suspected acute appendicitis: a prospective, randomized, controlled study comparing three different techniques. *Dis Colon Rectum* 2007;50(8):1223-9.
63. Lee CC, Golub R, Singer AJ, Cantu R, Levinson H. Routine versus selective abdominal computed tomography scan in the evaluation of right lower quadrant pain: a randomized controlled trial. *Acad Emerg Med* 2007;14(2):117-22.
64. Walker S, Haun W, Clark J, McMillin K, Zeren F, Gilliland T. The value of limited computed tomography with rectal contrast in the diagnosis of acute appendicitis. *Am J Surg* 2000;180(6):450-4.
65. Anderson SW, Soto JA, Lucey BC, Ozonoff A, Jordan JD, Ratevosian J, et al. Abdominal 64-MDCT for suspected appendicitis: the use of oral and IV contrast material versus IV contrast material only. *AJR Am J Roentgenol* 2009;193(5):1282-8.
66. Mittal VK, Goliath J, Sabir M, Patel R, Richards BF, Alkalay I, et al. Advantages of focused helical computed tomographic scanning with rectal contrast only vs triple contrast in the diagnosis of clinically uncertain acute appendicitis: a prospective randomized study. *Arch Surg* 2004;139(5):495-9.
67. Hong JJ, Cohn SM, Ekeh AP, Newman M, Salama M, Leblang SD, et al. A prospective randomized study of clinical assessment versus computed tomography for the diagnosis of acute appendicitis. *Surg Infect* 2003;4(3):231-9.
68. Khanal BR, Ansari MA, Pradhan S. Accuracy of ultrasonography in the diagnosis of acute appendicitis. *Kathmandu Univ Med J (KUMJ)* 2008;6(1):70-4.
69. Wilson EB, Cole JC, Nipper ML, Cooney DR, Smith RW. Computed tomography and ultrasonography in the diagnosis of appendicitis: when are they indicated? *Arch Surg* 2001;136(6):670-5.
70. Yu SH, Kim CB, Park JW, Kim MS, Radosevich DM. Ultrasonography in the diagnosis of appendicitis: evaluation by meta-analysis. *Korean J Radiol* 2005;6(4):267-77.
71. Krajewski S, Brown J, Phang PT, Raval M, Brown CJ. Impact of computed tomography of the abdomen on clinical outcomes in patients with acute right lower quadrant pain: a meta-analysis. *Can J Surg* 2011;54(1):43-53.
72. Al-Khayal KA, Al-Omran MA. Computed tomography and ultrasonography in the diagnosis of equivocal acute appendicitis. A meta-analysis. *Saudi Med J* 2007;28(2):173-80.
73. Anderson BA, Salem L, Flum DR. A systematic review of whether oral contrast is necessary for the computed tomography diagnosis of appendicitis in adults. *Am J Surg* 2005;190(3):474-8.
74. Weston AR, Jackson TJ, Blamey S. Diagnosis of appendicitis in adults by ultrasonography or computed tomography: a systematic review and meta-analysis. *Int J Technol Assess Health Care* 2005;21(3):368-79.
75. Obermaier R, Benz S, Asgharnia M, Kirchner R, Hopt UT. Value of ultrasound in the diagnosis of acute appendicitis: interesting aspects. *Eur J Med Res* 2003;8(10):451-6.
76. Doria AS. Optimizing the role of imaging in appendicitis. *Pediatr Radiol* 2009;39(Suppl 2):S144-S148.
77. Jacobs JE, Birnbaum BA, Macari M, Megibow AJ, Israel G, Maki DD, et al. Acute appendicitis: comparison of helical CT diagnosis focused technique with oral contrast material versus nonfocused technique with oral and intravenous contrast material. *Radiology* 2001;220(3):683-90.
78. Lane MJ, Liu DM, Huynh MD, Jeffrey RB, Mindelzun RE, Katz DS. Suspected acute appendicitis: nonenhanced helical CT in 300 consecutive patients. *Radiology* 1999;213(2):341-6.
79. Sauerland S, Agresta F, Bergamaschi R, Borzellino G, Budzynski A, Champault G, et al. Laparoscopy for abdominal emergencies: evidence-based guidelines of the European Association for Endoscopic Surgery. *Surg Endosc* 2006;20(1):14-29.
80. Society of American Gastrointestinal and Endoscopic Surgeons. Diagnostic laparoscopy guidelines. Los Angeles: SAGES; 2007. <http://www.sages.org/publication/id/12/>
81. Sauerland S, Jaschinski T, Neugebauer EA. Laparoscopic versus open surgery for suspected appendicitis. *Cochrane Database Syst Rev* 2010;10:CD001546.
82. Vettoretto N, Gobbi S, Corradi A, Belli F, Piccolo D, Pernazza G, et al. Consensus conference on laparoscopic appendectomy: development of guidelines. *Colorectal Dis* 2011;13(7):748-54.
83. van Dalen R, Bagshaw PF, Dobbs BR, Robertson GM, Lynch AC, Frizelle FA. The utility of laparoscopy in the diagnosis of acute appendicitis in women of reproductive age. *Surg Endosc* 2003;17(8):1311-3.
84. Larsson PG, Henriksson G, Olsson M, Boris J, Stroberg P, Tronstad SE, et al. Laparoscopy reduces unnecessary appendectomies and improves diagnosis in fertile women. A randomized study. *Surg Endosc* 2001;15(2):200-2.
85. Morino M, Pellegrino L, Castagna E, Farinella E, Mao P. Acute nonspecific abdominal pain: A randomized, controlled trial comparing early laparoscopy versus clinical observation. *Ann Surg* 2006;244(6):881-6.

86. Dominguez LC, Sanabria A, Vega V, Osorio C. Early laparoscopy for the evaluation of nonspecific abdominal pain: a critical appraisal of the evidence. *Surg Endosc* 2011;25(1):10-8.
87. Sauerland S, Lefering R, Neugebauer EA. Laparoscopic versus open surgery for suspected appendicitis. *Cochrane Database Syst Rev* 2004;4:CD001546.
88. Moberg AC, Berndsen F, Palmquist I, Petersson U, Resch T, Montgomery A. Randomized clinical trial of laparoscopic versus open appendicectomy for confirmed appendicitis. *Br J Surg* 2005;92(3):298-304.
89. Wilms IM, de Hoog DE, de Visser DC, Janzing HM. Appendectomy versus antibiotic treatment for acute appendicitis. *Cochrane Database Syst Rev* 2011;11:CD008359.
90. Varadhan KK, Neal KR, Lobo DN. Safety and efficacy of antibiotics compared with appendicectomy for treatment of uncomplicated acute appendicitis: meta-analysis of randomised controlled trials. *BMJ* 2012;344:e2156.
91. Kim K, Kim YH, Kim SY, Kim S, Lee YJ, Kim KP, et al. Low-dose abdominal CT for evaluating suspected appendicitis. *N Engl J Med* 2012;366(17):1596-605.
92. Société française de radiologie, Assistance publique hôpitaux de Paris., Société francophone de médecine d'urgence. Urgences de l'adulte : objectifs pour l'indication d'imagerie. Paris: AP-HP; 2004.
http://www.sfrnet.org/data/upload/files/a7e7222e420ac736c1256b6c0044cb07/14668_urgence_adulte.pdf
93. Drake FT, Florence MG, Johnson MG, Jurkovich GJ, Kwon S, Schmidt Z, et al. Progress in the diagnosis of appendicitis: A report from Washington State's surgical care and outcomes assessment program. *Ann Surg* 2012;256(4):586-94.
94. Poortman P, Oostvogel HJ, Bosma E, Lohle PN, Cuesta MA, de Lange-de Klerk ES, et al. Improving diagnosis of acute appendicitis: results of a diagnostic pathway with standard use of ultrasonography followed by selective use of CT. *J Am Coll Surg* 2009;208(3):434-41.
95. van Randen A., Lameris W, van Es HW, van Heesewijk HP, van Ramshorst B, Ten Hove W, et al. A comparison of the accuracy of ultrasound and computed tomography in common diagnoses causing acute abdominal pain. *Eur Radiol* 2011;21(7):1535-45.
96. Delacour H, François N, Servonnet A, Gentile A, Roche B. Les rapports de vraisemblance : un outil de choix pour l'interprétation des tests biologiques. *Immunoanal Biol Spec* 2009;24(2):92-9.
97. François Y, Bonvoisin S, Descos L, Vignal J. Etude prospective d'un score prédictif d'appendicite dans les douleurs de la fosse iliaque droite. Devenir des patients à long terme. *Gastroenterol Clin Biol* 1991;15(11):794-9.
98. Moberg AC, Ahlberg G, Leijonmarck CE, Montgomery A, Reiertsen O, Rosseland AR, et al. Diagnostic laparoscopy in 1043 patients with suspected acute appendicitis. *Eur J Surg* 1998;164(11):833-40.

Fiche descriptive

Intitulé	Descriptif
Méthode de travail	Rapport d'évaluation technologique
Date de mise en ligne	Novembre 2012
Date d'édition	Uniquement disponible sous format électronique sur www.has-sante.fr
Objectif(s)	Les objectifs tels que validés par le collège de la HAS le 20 juillet 2012 visent à produire les référentiels qui définissent la stratégie diagnostique permettant de poser l'indication d'une appendicectomie.
Professionnel(s) concerné(s)	Chirurgiens digestifs, gastro-entérologues, radiologues, urgentistes, biologistes, anatomopathologiste, médecins généralistes, gynécologues-obstétriciens
Demandeur	CNAMTS, DGOS, DSS
Promoteur	Haute autorité de santé (HAS), service évaluation des actes professionnels (SEAP)
Pilotage du projet	Coordination : Nadia ZEGHARI-SQUALLI, chef de projet, SEAP (chef de service : Sun-Hae LEE-ROBIN, adjoint au chef de service : Michèle MORIN-SURROCA) Secrétariat : Banedé SAKO et Christine MAYOL, Assistantes, SEAP
Participants	Expertise externe à la HAS : Groupe d'orientation : Hervé BERCHE, Stéphanie FRANCHI-ABELLA, Philippe LEVEAU, Philippe MARTEAU, Christophe PONCELET, Nathalie SIAUVE, Karem SLIM, Benoit TERRIS, Bertrand TOUSSAINT Groupe de cotation : Claude BENDAVID, Jacques BIENVENU, Marie-Jeanne BOUDET, André DABROWSKI, Pierre DARNAULT, François ESCAT, Eric FRAMPAS, Agnès LEURET, Jean-Marc LORPHELIN, Henri MARRET, Muriel MATHONNET, Bertrand MORIN, Jean-Marc PAULY, Jacques REMIZE, Janick SELVES, Jean SENDE, Magaly ZAPPA
Recherche documentaire	D'octobre 2011 à avril 2012 (stratégie de recherche documentaire décrite en annexe 3) Réalisée par Sophie DESPEYROUX, documentaliste, avec l'aide de Yasmine LOMBRY, assistante documentaliste, sous la responsabilité de Frédérique PAGES, chef du service documentation - information des publics, et Christine DEVAUD, adjointe au chef de service
Auteurs de l'argumentaire	Nadia ZEGHARI-SQUALLI, chef de projet, SEAP, sous la responsabilité de Michèle MORIN-SURROCA, adjoint au chef de service
Validation	Examen par la Commission d'évaluation économique et santé publique (CEESP) : octobre 2012 Collège de la HAS : novembre 2012
Autres formats	Pas d'autre format que le format électronique disponible sur www.has-sante.fr
Documents d'accompagnement	Note de cadrage, texte court, décision HAS (novembre 2012) disponibles sur www.has-sante.fr

N° ISBN : 978-2-11-128575-0



Toutes les publications de l'HAS sont téléchargeables sur
www.has-sante.fr