

Recommandations pour la pratique clinique

Révision méthodologique du processus d'élaboration des recommandations pour la pratique clinique

Élaborées par le Collège national des gynécologues
et obstétriciens français

Promoteur

CNGOF (Collège national des gynécologues et obstétriciens français)
91 boulevard de Sébastopol – 75002 Paris

Comité d'organisation

T. SCHMITZ, coordinateur (gynécologue-obstétricien, CHU Robert-Debré, Paris, CNGOF), L. SENTILHES (gynécologue-obstétricien, CHU, Bordeaux, CNGOF), M.V. SENAT (gynécologue-obstétricien, CHU Bicêtre, Le Kremlin-Bicêtre, CNGOF), X. FRITEL (gynécologue-obstétricien, CHU, Poitiers, CNGOF)

Experts du groupe de travail

É. AZRIA (gynécologue-obstétricien, ESPIC, Paris), C. DENEUX-THARAUX (épidémiologiste, Inserm), C. HUCHON (gynécologue-obstétricien, CH, Poissy), N. BOURDEL (gynécologue-obstétricien, CHU, Clermont-Ferrand), A. FAUCONNIER (gynécologue-obstétricien, CH, Poissy)

Lecteurs

J. BOUYER (épidémiologiste, Inserm), P. DERUELLE (gynécologue-obstétricien, CHU, Strasbourg), D. GALLOT (gynécologue-obstétricien, CHU, Clermont-Ferrand), C. GARABEDIAN (gynécologue-obstétricien, CHU, Lille), F. GOFFINET (gynécologue-obstétricien, CHU, Paris),

G. KAYEM (gynécologue-obstétricien, CHU, Paris), V. LAVOUÉ (gynécologue-obstétricien, CHU, Rennes), C. LE RAY (gynécologue-obstétricien, CHU, Paris), E. LORTHE (épidémiologiste, Inserm), N. SANANÈS (gynécologue-obstétricien, CHU, Strasbourg), C. VAYSSIÈRE (gynécologue-obstétricien, CHU, Toulouse)

Résumé

Objectifs - Réviser l'organisation et la méthodologie des recommandations pour la pratique clinique (RPC) du Collège national des gynécologues et obstétriciens français (CNGOF).

Méthodes - Les différentes méthodes d'organisation et de cotation de la preuve scientifique ont été consultées après recherche dans la base de données Medline.

Résultats - Le groupe RPC du CNGOF a décidé d'adopter une organisation des RPC basée sur les 23 critères AGREE II (pour Appraisal of Guidelines for REsearch and Evaluation) et de coter la preuve scientifique selon la méthode GRADE (pour Grading of Recommendation Assessment, Development, and Evaluation).

Conclusion - En adoptant les critères du consortium AGREE II et la cotation de la preuve selon la méthode GRADE, le CNGOF augmentera la qualité de l'ensemble du processus, délivrera des messages plus ciblés et facilement assimilables, pour faciliter mieux encore les prises de décisions des professionnels, au bénéfice espéré de la santé des femmes et de leurs enfants.

Mots clés : recommandations pour la pratique clinique, système GRADE, consortium AGREE II, question PICO

INTRODUCTION

Les premières recommandations pour la pratique clinique (RPC) élaborées par le Collège national des gynécologues et obstétriciens français (CNGOF) portant sur la prise en charge des infections cervico-vaginales pendant la grossesse ont été publiées en 1997, il y a déjà plus de 20 ans [1]. Elles constituaient à l'époque l'une des premières tentatives françaises visant à intégrer méthodiquement la médecine basée sur les preuves dans la pratique clinique au sein de notre spécialité, afin de faciliter et d'homogénéiser les prises de décisions par les différents acteurs du système de soins, et finalement d'améliorer la prise en charge des patientes. Répondant à une attente des soignants de la gynécologie, de l'obstétrique et plus largement de la périnatalogie dans son ensemble, l'intérêt a été rapidement au rendez-vous, et sans cesse croissant, comme en témoigne l'importante fréquentation des sessions RPC aux journées nationales du CNGOF. Cependant, il est apparu, depuis plusieurs années déjà, que le système d'organisation et de rédaction de ces RPC avait probablement atteint ses limites, voire était désormais dépassé, en particulier dans l'exercice consistant à traduire le niveau de la preuve scientifique, c'est-à-dire basé sur la synthèse des résultats des études publiées (le niveau de preuve) en force de recommandation, qui nécessite d'intégrer une dimension de faisabilité et de rapport bénéfice/risque en vie réelle (le grade). En effet, parallèlement aux progrès de la médecine fondée sur les preuves ces 25 dernières années, se sont développées de nouvelles techniques de gradation de la preuve scientifique ainsi que de nouvelles approches méthodologiques, des processus systématisés, pour élaborer des RPC, répondant à des cahiers des charges de plus en plus détaillés et exigeants. Force est de constater que la méthodologie encore en vigueur actuellement au CNGOF pour l'organisation des RPC ne répond plus de manière satisfaisante aux critères actuels de qualité. L'amélioration de la qualité méthodologique devrait aboutir à une meilleure assimilation des messages clés des RPC et réduire les écarts existants entre recommandations et pratiques cliniques constatées sur le terrain [2, 3].

Le groupe des RPC, aussi bien les gynécologues que les obstétriciens, a donc souhaité réviser la méthodologie des RPC du CNGOF. Le groupe s'est réuni 4 journées entre octobre 2018 et mai 2019 afin de définir la façon dont elles seraient désormais organisées. Le texte présenté ici résume les décisions qui ont été prises lors de ces réunions et l'organisation future des RPC.

I. NOUVELLE ORGANISATION DES RPC DU CNGOF

I.a. Détermination des thématiques des RPC

Jusqu'à présent, le processus aboutissant au choix des thématiques des RPC était mal systématisé. Comme cela a été le cas jusqu'à présent, le conseil scientifique (CS) et le groupe RPC proposeront et valideront les thématiques des RPC. Cependant, une saisine du CS par les associations d'utilisateurs, les agences gouvernementales (ARS, HAS, etc.), et organismes divers (Santé publique France, Gynerisq, etc.) sera possible afin qu'ils proposent les sujets qu'ils leur sembleraient importants de traiter. Le CS hiérarchisera ces demandes, déterminera si elles relèvent d'une RPC, et dans les cas contraires, proposera alors la méthodologie la plus adéquate pour y répondre.

I.b. Critères AGREE II

Le groupe de travail a considéré que les RPC du CNGOF devaient désormais suivre les critères du consortium AGREE II (pour *Appraisal of Guidelines for REsearch & Evaluation*) [4]. Le consortium AGREE II est un groupe international de cliniciens, chercheurs et développeurs de RPC ayant élaboré une grille définissant les critères de qualité d'une RPC. Cette grille a été publiée pour la première fois en 2003 et dans une version révisée en 2013 (Tableau 1). Elle fait actuellement référence comme outil d'évaluation de la qualité d'une RPC et est utilisée à ce titre lorsque sont pratiquées des comparaisons internationales de RPC sur une même thématique [5].

Même si la majorité des critères de la grille AGREE II était respectée ces dernières années, des manquements dans la rédaction des RPC faisaient apparaître que certains critères étaient souvent incomplètement remplis ou mal formalisés, laissant croire, parfois à tort, à une qualité dégradée des RPC du CNGOF. Le non-respect des critères AGREE II a conduit par le passé à l'exclusion de nos RPC de certaines comparaisons internationales de recommandations [6] ou à des appréciations péjoratives [7]. L'adoption par le CNGOF des critères qualité AGREE II (Tableau 1) devrait permettre à l'avenir que nos RPC soient facilement identifiées comme répondant aux standards actuels de qualité. La grille AGREE II comprend 23 critères de qualité, répartis en 6 domaines, qui sont repris ci-dessous.

Tableau 1 - Grille d'évaluation de la qualité des recommandations pour la pratique clinique (AGREE II)

Grille d'évaluation de la qualité des recommandations pour la pratique clinique (AGREE II)
Domaine 1. Champ et Objectifs
1. Le ou les objectifs de la RPC sont décrits explicitement
2. La ou les questions cliniques couvertes par la RPC sont décrites explicitement
3. La population (patients, public, etc.) à laquelle la RPC doit s'appliquer est décrite explicitement
Domaine 2. Participation des groupes concernés
4. Le groupe ayant élaboré la RPC inclut des représentants de tous les groupes professionnels concernés
5. Les opinions et les préférences de la population cible (patients, public, etc.) ont été identifiés
6. Les utilisateurs cibles de la RPC sont clairement définis
Domaine 3. Rigueur d'élaboration de la RPC
7. Des méthodes systématiques ont été utilisées pour rechercher les preuves scientifiques
8. Les critères de sélections des preuves sont clairement décrits
9. Les forces et les limites des preuves scientifiques sont clairement décrites
10. Les méthodes utilisées pour formuler les recommandations sont clairement décrites
11. Les bénéfices, les effets secondaires et les risques en termes de santé ont été pris en considération dans la formulation des recommandations
12. Il y a un lien explicite entre les recommandations et les preuves scientifiques sur lesquelles elles reposent
13. La RPC a été revue par des experts externes avant sa publication
14. Une procédure d'actualisation de la RPC est décrite
Domaine 4. Clarté et présentation
15. Les recommandations sont précises et sans ambiguïté
16. Les différentes options de prise en charge de l'état ou du problème de santé sont clairement présentées
17. Les recommandations clés sont facilement identifiables
Domaine 5. Applicabilité
18. La RPC décrit les éléments facilitant son application et les obstacles
19. La RPC offre des conseils et/ou des outils pour mettre les recommandations en pratique
20. Les répercussions potentielles sur les ressources de l'application des recommandations ont été examinées
21. La RPC propose des critères de suivi et de vérification
Domaine 6. Indépendance éditoriale
22. Le point de vue des organismes de financement n'a pas influencé le contenu de la RPC
23. Les intérêts divergents (les conflits d'intérêts) des membres du groupe ayant élaboré la RPC ont été pris en charge et documentés

I.b.1. Domaine 1 : champs et objectifs (critères 1 à 3)

Les 3 premiers critères de la grille AGREE II portent sur les objectifs (critère 1), les questions cliniques (critère 2) et la population cible (critère 3) de la RPC. Si ces 3 critères étaient le plus souvent clairement énoncés et identifiables dans les textes courts publiés en français et en anglais, les questions cliniques n'étaient souvent que vaguement formulées initialement, ce qui les rendaient a posteriori inappropriées. En effet, les textes courts faisaient la synthèse des textes longs qui correspondaient chacun à une question, énoncée à la manière de ce qui est fait dans un chapitre d'ouvrage de référence, de façon insuffisamment ciblée (traitement de la menace d'accouchement prématuré, par exemple). Chacune de ces questions était ensuite découpée, là encore, comme des sous-chapitres d'un ouvrage de

référence. Ce découpage des questions, s'il avait le mérite de rendre les textes longs très exhaustifs, ne permettait souvent pas d'aboutir à des recommandations basées sur des preuves scientifiques concordantes et conduisait fréquemment à l'émission de recommandations sous la forme d'accord professionnel, qui ne reflétaient en fait que l'opinion des membres du groupe de travail et n'avaient de ce fait pas la légitimité des consensus d'experts obtenus par des méthodes plus structurées dont c'est l'objectif. La multiplication des recommandations basées sur des accords professionnels avait pour conséquence de noyer les recommandations importantes de haut grade et de les rendre difficilement identifiables, et finalement probablement moins bien assimilables par le public concerné. Une recommandation pour être pertinente doit non seulement être en accord avec les données scientifiques disponibles et valides, mais également satisfaire à l'exigence d'utilité en pratique courante. Celle-ci doit avoir un impact sur les pratiques pour améliorer réellement la qualité des soins. De ce point de vue, la formulation initiale des questions posées est importante. C'est pourquoi le groupe de travail a décidé que les questions seraient désormais formulées au format PICO (Tableau 2), comme recommandé par l'Organisation mondiale de la santé (OMS) [8].

Formulation des questions au format PICO

PICO est un acronyme signifiant *Population, Intervention, Comparison, Outcome*. La formulation des questions au format PICO constituera le premier changement méthodologique majeur des RPC futures. À chaque question PICO, correspondra une recommandation si la littérature existante présente un niveau de preuve suffisant pour y répondre.

Comme son nom l'indique, la formulation de chaque question au format PICO permet de faire apparaître dans la question : la population, l'intervention, la comparaison (ou comparateur) et le critère de jugement sur lesquels portera la littérature scientifique permettant d'émettre une recommandation pour une question donnée. La formulation de ces questions constitue la première étape de l'élaboration d'une RPC.

Ce mode de formulation de la question permet une analyse de la littérature plus ciblée et fermée, et évite ainsi des questions trop ouvertes aboutissent, en l'absence de données scientifiques, à des recommandations de type accord professionnel. Des exemples de questions au format PICO sont donnés (Tableau 2), cadrant la problématique du traitement de la menace d'accouchement prématuré (MAP), cité précédemment. Dans ces exemples, la population (P) à

laquelle va s'appliquer l'évaluation de la preuve scientifique est d'emblée ciblée (les femmes ayant une MAP) ou très ciblée (risque d'accouchement prématuré entre 34 et 37 SA), ainsi que l'intervention (I) (nifédipine) et le comparateur (C) (atosiban) pour un critère de jugement donné (O) (morbidité néonatale). Les questions PICO sont aussi adaptées pour formuler des questions sur la pertinence de tests diagnostiques, l'intervention (I) est alors la réalisation du test diagnostique et le critère de jugement (O), la maladie à diagnostiquer. Un exemple est donné dans le Tableau 2.

Tableau 2 - Formulation des questions au format PICO

Les questions PICO	
P (Population)	Population
I (Intervention)	Intervention
C (Comparison)	Comparaison
O (Outcome)	Critère de jugement
Exemples	
Chez les femmes en menace d'accouchement prématuré (P), la nifédipine (I), en comparaison à l'atosiban (C), est-elle plus efficace pour réduire la morbidité néonatale (O) ?	
Chez les femmes à risque d'accouchement prématuré entre 34 et 37 SA (P), l'administration de corticoïdes anténatals (I) réduit-elle la mortalité néonatale (O) ? (L'absence de comparateur sous-entend une comparaison à un placebo ou à l'absence de traitement)	
Chez les femmes présentant des douleurs pelviennes chroniques (P), l'IRM (I), en comparaison à l'échographie (C), est-elle plus performante pour diagnostiquer l'endométriiose profonde (O) ?	

La formulation d'une question au format PICO implique une classification de l'importance des critères de jugement, c'est-à-dire le O (*Outcome*) de PICO. Le but d'une recommandation étant de conduire à des bénéfices en terme de santé, ces bénéfices seront d'autant plus marqués qu'ils porteront sur des critères de jugement importants. L'importance des critères de jugement est portée sur une échelle allant de 1 (importance la plus faible) à 9 (importance cruciale) et est classée en trois niveaux (Tableau 3). De manière caricaturale et à titre d'exemple, pour une question portant sur le traitement de la menace d'accouchement prématuré, un critère de jugement comme le décès néonatal serait coté 9, la morbidité sévère cotée 8, la morbidité modérée 7, le terme d'accouchement avant 32 semaines d'aménorrhée (SA) 6, avant 34 SA 5, avant 37 SA 4. Le choix de ces critères ainsi que leurs classifications par degré d'importance doivent être réalisés a

priori, avant que ne soit débutée la recherche bibliographique approfondie. L'ensemble des membres du groupe de travail doit impérativement y participer, coordonnateur, méthodologiste, rédacteurs, mais aussi et en particulier, les représentants des usagers, comme nous le verrons plus loin.

Tableau 3 - Classification de l'importance des critères de jugement des questions au PICO

9	Cruciaux pour la prise de décision
8	
7	
6	Importants mais pas cruciaux pour la prise de décision
5	
4	
3	Pas importants
2	
1	

Enfin, il est prévu de limiter le nombre de critères de jugement, le choix de 2 critères de jugements par niveau d'importance (crucial / important mais pas crucial / pas important) apparaît comme un seuil à ne pas dépasser [8]. En effet, afin de pouvoir coter le niveau de la preuve scientifique apportée par la littérature, il faut que les critères de jugement choisis soient communs aux différentes publications et répondent à la question posée par le groupe de travail pour qu'une synthèse puisse être réalisée. La multiplication des critères de jugement réduirait la probabilité que ceux-ci aient été étudiés dans chaque essai ou étude et rendrait ainsi difficile, voire impossible, le rendu d'un niveau de preuve pour une question donnée. Le choix de ces critères de jugement devra, lorsque c'est possible, être en adéquation avec les critères de jugement ayant fait l'objet d'accords internationaux [9, 10].

1.b.2. Domaine 2 : participation des groupes concernés (critères 4 à 6)

Les critères 4 à 6 de la grille AGREE II portent sur la constitution du groupe participant aux RPC ainsi que sur l'identification des besoins des usagers.

Constitution du groupe de travail

Le groupe chargé de la révision de la méthodologie des RPC du CNGOF en accord avec le bureau du conseil d'administration (CA) et le CS a souhaité mieux formaliser la procédure de choix et les rôles du

président, du coordonnateur, du méthodologiste ainsi que des experts des RPC.

(i) Le président d'une RPC est choisi par le bureau du CA. Reconnu par ses pairs dans le domaine et incarnant la thématique, il est au fait et adhère à la méthodologie des RPC. Il anime les débats de façon non partisane ce qui implique absolument qu'il soit libre de conflits d'intérêt sur le sujet, et arbitre et gère les désaccords. Représentant le bureau du CA au sein de la RPC, il endosse la responsabilité morale et scientifique de la RPC au nom du collège.

(ii) Le coordonnateur est choisi par le CS. Compétent dans la thématique choisie, il est responsable du calendrier de la RPC, de la bonne organisation thématique du travail, de la rédaction du plan et du texte final. Il rédige les questions PICO qui seront proposées aux experts rédacteurs en lien avec le méthodologiste. Il fait le lien entre les membres du groupe, relance les experts, établit les comptes rendus de séance.

(iii) Le méthodologiste est choisi par le CS. Il est responsable de la qualité méthodologique de la RPC. Pour cela, il gère les conflits d'intérêts des membres du groupe en lien avec le CS, pilote la recherche bibliographique en lien avec le chargé de recherche documentaire et les experts rédacteurs auxquels il fournit une aide méthodologique, il guide les débats méthodologiques de façon non partisane et arbitre les conflits méthodologiques. Il rédige les questions PICO qui seront proposées aux experts rédacteurs en lien avec le coordonnateur.

(iv) Enfin, les experts rédacteurs seront choisis conjointement par le président, le coordinateur et le méthodologiste. Ils corrigent et valident les questions PICO proposées par le coordinateur et le méthodologiste, et rédigent les argumentaires et recommandations répondant à ces questions. Le nombre de deux experts par question PICO est proposé. Chaque expert peut répondre à plusieurs questions PICO. Les experts seront aidés dans leur recherche bibliographique par le chargé de recherche documentaire (qui sera idéalement un interne DES (diplôme d'études spécialisées) ou un CCA (chef de clinique assistant) ayant une formation complémentaire en épidémiologie et/ou recherche clinique). Le rôle du chargé de recherche documentaire dans la nouvelle organisation des RPC est majeur puisque c'est lui, comme nous le verrons plus loin, qui proposera la première cotation de la qualité de la preuve scientifique.

Participation des usagers

L'autre point important du domaine 2 est la place faite aux usagers. Si depuis de nombreuses années, des représentants des associations de patientes ont participé aux RPC, ils étaient souvent associés trop tardivement dans le processus d'élaboration des recommandations ce qui limitait leur implication. À l'avenir, ces représentants proposeront et participeront au choix des critères de jugement des questions PICO et jugeront de leur pertinence selon leur point de vue. Les associations reconnues et si possible agréées pour le domaine concerné seront désignées conjointement par le bureau du CA et le CS. Celles-ci proposeront leurs représentants. La participation de patients volontaires non issus d'associations est également envisageable pour certaines RPC.

I.b.3. Domaine 3 : rigueur d'élaboration de la RPC (critères 7 à 14)

Ce domaine est le plus volumineux de la grille AGREE II et celui qui va générer les changements les plus importants dans la nouvelle organisation proposée.

En effet, si les critères 7 et 8 étaient bien respectés et conformes aux exigences d'AGREE II, il est apparu que l'évaluation du niveau de la preuve scientifique et sa translation en recommandation (critères 9 à 12) par l'utilisation de la méthode proposée encore aujourd'hui par la Haute Autorité de santé (HAS), inchangée depuis plus de 20 ans puisque identique à celle que proposait l'ANDEM (Agence nationale pour le développement de l'évaluation médicale, 1991-1996) et l'ANAES (Agence nationale de l'accréditation et de l'évaluation en santé, 1996-2004), n'étaient plus en phase avec les exigences internationales actuelles et devaient être revues. Il est ainsi nécessaire d'adopter d'autres méthodes de gradation de la preuve scientifique.

Le groupe chargé de la révision de la méthodologie, après avoir examiné les différentes méthodes disponibles, a choisi d'adopter la méthode GRADE (pour *Grading of Recommendation Assessment, Development, and Evaluation*) [11-13], largement adoptée par un grand nombre de sociétés savantes internationales de diverses spécialités.

La gradation du niveau de la preuve scientifique par la méthode GRADE

Il s'agit, après la formulation des questions au format PICO, de la seconde modification majeure dans la méthodologie des RPC du CNGOF.

L'utilisation du système proposé par la HAS (Tableau 4) et utilisé jusqu'à présent par le CNGOF était rigide à double titre [14], pour la cotation du niveau de la preuve scientifique d'une part et pour sa

translation en force de recommandation, d'autre part. Enfin, ce système, exclusivement tourné vers la thérapeutique ne permettait pas de faire des recommandations de manière pertinente concernant les méthodes diagnostiques.

Tableau 4 - Gradation de la preuve et de la recommandation selon la méthode HAS

Niveau de preuve scientifique fourni par la littérature (études thérapeutiques)	Grade des recommandations
<p><i>Niveau 1</i></p> Essais comparatifs randomisés de forte puissance Méta-analyses d'essais comparatifs randomisés Analyse de décision basée sur des études bien menées	A Preuve scientifique établie
<p><i>Niveau 2</i></p> Essais comparatifs randomisés de faible puissance Études comparatives non randomisées bien menées Études de cohorte	B Présomption scientifique
<p><i>Niveau 3</i></p> Études cas-témoins	C Faible niveau de preuve
<p><i>Niveau 4</i></p> Études comparatives comportant des biais importants Études rétrospectives Séries de cas	
<p><i>Niveau d'experts</i></p>	Accord professionnel

Du fait d'une corrélation absolue entre la conception méthodologique de l'étude et le niveau de preuve, un essai randomisé multicentrique conduisait presque systématiquement à un niveau de preuve maximal, quand bien même la validité interne et/ou externe de l'étude aurait été jugée insuffisante ou le critère de jugement ne couvrant que partiellement l'objectif de la recommandation. Inversement, plusieurs études de cohortes prospectives, bien conduites, avec des effectifs importants, des résultats concordants, prenant en compte statistiquement les biais potentiels, apportaient le plus souvent un niveau de preuve très inférieur à celui des essais randomisés.

Surtout, la force de la recommandation était obligatoirement corrélée au niveau de la preuve scientifique, c'est-à-dire, pour reprendre un exemple caricatural devenu désormais célèbre [15] qu'il était impossible, en l'absence d'essai randomisé, de recommander avec un grade élevé, l'utilisation d'un parachute pour sauter d'un avion. Dans cet exemple, c'est l'aspect éthique de la recommandation qui n'est pas pris en compte. D'autres exemples concernent la faisabilité en vie réelle, par exemple l'essai CARE [16] qui propose pour la prophylaxie de l'incontinence post-opératoire de la promonto-fixation

par laparotomie (actuellement remplacé par la coéloscopie) l'utilisation systématique d'une intervention abandonnée (la colpo-suspension selon Burch).

Le système GRADE permet de s'affranchir de la plupart des défauts du système HAS, mais est de maniement plus complexe et nécessitera un effort de pédagogie afin que les cliniciens, à qui sont destinés les RPC, l'assimilent et se l'approprient pleinement. Cet effort pourra passer par l'organisation d'atelier et de cours de méthodologie sur Internet. Le système GRADE a lui aussi été développé par un consortium international. Initié en l'an 2000, par un groupe de chercheurs principalement canadiens, norvégiens, italiens, américains et suisses, ce système a été conçu pour combler les lacunes des systèmes de classement dédiés à l'évaluation de la preuve scientifique existants alors.

La grille de cotation de la preuve scientifique selon la méthode GRADE est donnée dans le Tableau 5. Ses principes fondamentaux sont les suivants.

(i) La première caractéristique majeure différenciant le système GRADE du système HAS est de nature sémantique. En effet, on ne fait plus référence ici à un niveau de preuve (*Level of evidence*) mais à la qualité de la preuve (*Quality of evidence*). On passe donc d'une analyse en apparence purement quantitative de la preuve à une analyse comportant une importante dimension qualitative, plus en phase avec la réalité de l'évaluation et qui permet par exemple de tenir compte des biais, de la généralisabilité, et du caractère universel ou non de la preuve.

(ii) Il ne s'agit plus de discuter du niveau de preuve de chaque essai ou de chaque étude pris indépendamment, mais de la synthèse de l'ensemble de la littérature répondant à une question PICO. Le niveau de la preuve sera ainsi donné pour estimer l'ensemble de la littérature scientifique évaluant par exemple un traitement I par rapport à un traitement C dans une population P pour un critère de jugement donné O. L'attribution d'un niveau de preuve pour le critère de jugement d'une question PICO découle bien entendu d'un travail de synthèse de l'ensemble de la littérature sur le sujet. Ce travail de synthèse est résumé dans des tableaux GRADE reprenant l'ensemble de la littérature, chaque étude y occupant une ligne. Ces tableaux seront réalisés par les chargés de recherche documentaire.

(iii) Il n'existe plus que 2 niveaux de preuve de base avant analyse : un niveau élevé (+4), attribué uniquement aux essais randomisés, et un niveau bas (+2) attribué aux études observationnelles (colonnes 1 et 2, Tableau 5).

Tableau 5 - Gradation de la preuve scientifique selon la méthode GRADE

1	2	3	4	5	6
Niveau initial de confiance ou de qualité de la preuve	Niveau initial de confiance dans l'estimation de l'effet ou de qualité de la preuve	Diminution ou majoration du niveau initial de confiance ou de qualité de la preuve Diminution - Risque de biais (-1) - Inconsistance (-1) - Effet indirect (-1 à -2) - Imprécision (-1) - Biais de publication (-1)	Augmentation - Forte amplitude de l'effet (+1 à +2) - Effet dose (+1) - Prise en compte des facteurs de confusion (+1)	Niveau final de confiance en la preuve	Niveau final de qualité de la preuve
	Méthodologie de l'étude				
Essai contrôlé randomisé	Élevée (+4)			Des études futures ont peu de chance de modifier notre confiance dans l'estimation de l'effet	Elevé +4
				Des études futures auront probablement un impact important sur notre confiance dans l'estimation de l'effet et pourraient modifier l'estimation	Moderé +3
Étude observationnelle	Basse (+2)			Des études futures auront très probablement un impact important sur notre confiance dans l'estimation de l'effet et modifieront probablement l'estimation	Bas +2
				L'estimation de l'effet est très incertaine	Très bas +1

(iv) Contrairement au système HAS, il est prévu par le système GRADE de pouvoir diminuer ou augmenter le niveau de preuve initial, et ainsi de faire entrer dans l'analyse la dimension qualitative. Le niveau de preuve élevé attribué à la littérature montrant, au moyen d'essais randomisés, la supériorité du traitement I par rapport au traitement C dans la population P pour le critère de jugement O sera abaissé si ces essais sont de qualité discutable (-1) ou mauvaise (-2), si leurs résultats sont discordants (-1), ou aboutissant à une estimation imprécise de l'effet du traitement C (-1), si l'effet du traitement C est indirect (-1 à -2) ou s'il existe des biais de publication (-1) (colonne 3, Tableau 5). À l'inverse, le niveau de preuve de l'effet d'un traitement I' par rapport à un traitement C' dans une population P' pour le critère de jugement O' déterminé à partir d'études observationnelles peut passer de bas (+2) à modéré (+3) ou même élevé (+4) si plusieurs des critères suivants sont remplis : preuve d'une association forte (risque relatif (RR) > 2 ou < 0.5) (+1) ou très forte (RR > 5 ou < 0.2) (+2), mise en évidence d'un effet dose (+1), prise en compte des facteurs de confusion connus (+1) (colonne 4, Tableau 5). Le niveau de preuve donné par la littérature répondant à une question PICO peut finalement varier de très bas (+1) à élevé (+4) (colonne 6, Tableau 5).

(v) Enfin, ces niveaux de preuve traduisent la confiance que nous avons dans l'effet du traitement I par rapport au traitement C dans la population P pour le critère de jugement O (colonne 5, Tableau 5). Il faut comprendre la confiance dans l'estimation comme la largeur de l'intervalle de confiance à 95 % et l'estimation de l'effet comme la valeur du risque relatif.

Maintenant qu'ont été expliqués les principes sur lesquels reposent la cotation de la preuve dans le système GRADE, nous ne ferons désormais plus référence dans cet article, au niveau de preuve mais uniquement à la qualité de la preuve pour ce système de cotation.

Formulation d'une recommandation selon la méthode GRADE

Alors que la formulation d'une recommandation selon la méthode HAS n'était que la translation d'un niveau de preuve d'une échelle de 1 à 4 en une recommandation d'un grade A à C et la correspondance de l'avis d'expert en accord professionnel, l'utilisation de la méthode GRADE fait entrer d'autres dimensions dans l'élaboration de la recommandation (Tableau 6). Si la qualité de la preuve reste l'élément prédominant, interviennent aussi l'importance de la pathologie traitée (incidence et sévérité), les valeurs et les préférences des soignants et des patients, la balance bénéfiques/risques, l'acceptabilité et la faisabilité du traitement ou d'une stratégie de soins, son coût, le risque d'iniquité en

fonction des populations. Le système GRADE ne fournit pas de méthode quantitative pour pondérer et finalement intégrer ces différentes dimensions dans une recommandation, car cette intégration dépend des valeurs et des ressources du pays et des populations auxquelles s'appliquera la RPC.

Tableau 6 - Dimensions intégrées dans la formulation d'une recommandation selon la méthode GRADE

Dans chaque recommandation, sont intégrées les dimensions suivantes :
<ul style="list-style-type: none"> - La qualité de la preuve scientifique mais aussi, • L'importance de la pathologie • Les valeurs et les préférences des soignants et des patientes • La balance bénéfiques/risques • L'acceptabilité et la faisabilité • Le coût • L'équité sociale

Si dans la majorité des cas, à une qualité élevée de la preuve correspondra une recommandation forte, et inversement, il est cependant possible avec le système GRADE de dissocier la qualité de la preuve de la force de la recommandation, suite à l'intégration des autres dimensions participant à son l'élaboration [17, 18]. Il est alors utile de justifier cette décision. Cinq situations ont été identifiées comme pouvant conduire à une recommandation forte malgré une qualité de la preuve basse ou très basse [17, 18]. Il s'agit i) des situations de menace vitale ou catastrophique (exemple du saut en parachute) (recommandation forte pour), ii) d'un bénéfice incertain associé à un mal certain, iii) de bénéfices potentiellement équivalents pour deux options thérapeutiques mais dont l'une est clairement moins risquée ou moins coûteuse que l'autre, iv) des bénéfices clairement équivalents pour deux options thérapeutiques mais dont l'une est potentiellement moins risquée ou moins coûteuse que l'autre, et enfin, v) des traitements potentiellement associés à des effets catastrophiques (recommandation forte contre).

Au final, l'intégration des différents paramètres évoqués précédemment aboutit à des recommandations qui ne comprennent finalement que deux dimensions : un sens, positif ou négatif (il est recommandé de..., il est recommandé de ne pas...), et une force, (faible ou forte). Le sens de la recommandation sera donné dans l'intitulé de la recommandation suivi de la qualité de la preuve scientifique et de la force de la recommandation entre parenthèses comme dans l'exemple ci-dessous :

Il est recommandé d'administrer des corticoïdes anténatals aux femmes à risque d'accouchement prématuré avant 34 SA (qualité de la preuve élevée, recommandation forte).

Les questions PICO sont formulées a priori, avant l'analyse de la littérature. Par conséquent, il arrivera que certaines questions ne trouvent pas de réponses lorsque la littérature scientifique sera insuffisante pour répondre à la question. La liste des questions auxquelles il n'aura pas été possible de répondre sera donnée en fin de RPC, précédée de la mention suivante : « Les données de la littérature scientifique sont actuellement insuffisantes pour répondre aux questions suivantes ».

Le point 13 du domaine 3 concerne la validation externe des recommandations par un groupe d'experts indépendant du groupe les ayant rédigées. Les recommandations du CNGOF continueront d'être envoyées à un panel de relecteurs. Il sera demandé à chaque relecteur de coter son niveau d'accord avec la recommandation, à la fois sur le fond mais aussi sur la forme. Les recommandations qui ne rencontreront pas un niveau d'accord suffisant seront reformulées ou même abandonnées par le groupe de travail.

Enfin, les RPC devront être réactualisées (point 14). Aucune procédure n'était proposée jusqu'à présent, mais il va de soi qu'une RPC tend à devenir plus ou moins rapidement obsolète en fonction de l'apparition de nouvelles connaissances sur la pathologie. La procédure suivante sera suivie. Le CS assurera une veille scientifique et fera le point sur la validité persistante ou non des données ayant permis la rédaction de la RPC. Il décidera de la nécessité d'une révision et du champ de celle-ci, soit après saisie du bureau du CA par des usagers ou des acteurs de santé, soit après alerte du coordonnateur et du méthodologiste à la lumière des avancées scientifiques dans un domaine donné. Le coordonnateur, le méthodologiste et le ou les experts des questions concernées par la révision de la RPC rédigeront un nouvel argumentaire pour la ou les questions PICO. En cas de nouvelle question, le coordonnateur proposera de nouveaux experts au groupe initial de la RPC.

1.b.4. Domaine 4 : clarté et présentation (critères 15 à 17)

Le domaine 4 comprend 3 critères portant sur la précision, la clarté et l'absence d'ambiguïté des recommandations. Le passage au système GRADE avec des questions formulées au format PICO devrait permettre de nettes améliorations dans ce domaine, en

réduisant, d'une part, le nombre total de recommandations et rendant ainsi plus visibles les recommandations fortes, et en faisant disparaître, d'autre part, le jargon existant dans lequel il fallait comprendre « qu'il n'est pas recommandé » signifiait que la littérature était insuffisante pour recommander l'option discutée mais que son adoption en pratique clinique était permise.

I.b.5. Domaine 5 : applicabilité (critères 18 à 21)

Le domaine 5 porte sur la diffusion et l'applicabilité des RPC.

Les RPC seront présentées en exclusivité aux Journées Paris Santé Femmes et diffusées par la suite lors de congrès régionaux. Elles feront l'objet de formation dans le cadre du développement professionnel continu (DPC). Elles sont la propriété intellectuelle, des membres du groupe du travail, du CNGOF qui les rendra accessibles en ligne sur le site du Collège, et des autres sociétés savantes qui auraient participé. Elles feront l'objet de publications en langue française et anglaise.

La RPC proposera des critères de suivi de son application en identifiant les recommandations fortes devant en être l'objet. Ces critères pourront par exemple devenir des indicateurs de qualité de prise en charge pour les réseaux périnataux. Le groupe des RPC n'est pas responsable de la mise en œuvre du suivi de ces indicateurs.

Enfin, le groupe pourra proposer les questions de recherche qui lui apparaissent comme prioritaires sur la thématique de la RPC, en cas de manque dans la littérature, en particulier pour les questions PICO intégrant des critères de jugement cruciaux qui n'auraient pas trouvé de réponse.

I.b.6. Domaine 6 : indépendance éditoriale (critères 22 et 23)

Le domaine 6 concerne l'indépendance éditoriale dont a pu bénéficier le groupe de travail lors de la rédaction des RPC. Le point 22 a trait aux organismes financeurs. Le CNGOF et son CS continueront à assurer toute l'indépendance nécessaire au groupe de travail comme cela a été le cas par le passé.

Le point 23 concerne les individus participants au groupe de travail. Si par le passé, une DPI (déclaration publique d'Intérêt) leur était demandée, celle-ci n'était parfois remise qu'à la toute fin du processus rédactionnel, et des conflits d'intérêts ont parfois été mis en évidence alors que tout retour en arrière était impossible. Il a été décidé qu'un groupe de travail ne pourrait être considéré comme constitué que lorsque l'ensemble des DPI de tous les membres participants aura été recueilli. Les DPI seront sur le modèle de celles déjà existantes à la HAS. Le CS nommera 2 rapporteurs qui étudieront

les DPI du président, du coordonnateur et du méthodologiste. Les DPI des experts rédacteurs et des représentants des usagers seront examinées par le méthodologiste et le coordonnateur. Les conflits d'intérêt portés à la connaissance du CS seront remontés au CA qui tranchera. La secrétaire administrative du CNGOF archivera les DPI, qui seront des documents juridiquement opposables dans le futur.

II. CONSÉQUENCES PRATIQUES DE LA RÉVISION DE LA MÉTHODOLOGIE DES RPC DU CNGOF

La modification de la méthodologie des RPC du CNGOF va avoir plusieurs conséquences importantes qui marqueront un véritable changement en comparaison avec ce qui a été fait jusqu'à présent. Elles tiennent presque toutes aux modalités de formulation des questions sous le format PICO.

II.a. Réduction du nombre de réunions

La première conséquence importante du changement de méthodologie sera une réduction du nombre de réunions. Une première réunion impliquera le président, le méthodologiste, le coordonnateur. Ils rédigeront ensemble une première version des questions PICO et choisiront le reste du groupe de travail. La deuxième réunion concernera l'ensemble du groupe de travail, et seront donc aussi présents, les experts rédacteurs ainsi que les représentants des usagers. Les questions PICO seront corrigées. Ces deux réunions seront assez rapprochées. À la suite de cette réunion, le/les chargés de recherche documentaire adresseront à l'ensemble du groupe de travail, les tableaux synthétiques évaluant la qualité de la preuve scientifique selon la méthode GRADE. Les experts rédigeront leurs argumentaires à partir de ces tableaux et proposeront leurs recommandations. Les troisième et quatrième réunions seront dévolues à la correction des argumentaires et des recommandations par le groupe de travail avant leur envoi aux relecteurs. La cinquième et dernière réunion consistera en l'intégration des modifications proposées par les relecteurs. Au total, une RPC ne devrait plus comprendre que 5 réunions au lieu de 9 actuellement pour l'obstétrique et de 6 à 8 pour la gynécologie.

II.b. Réduction du nombre de recommandations avec accords professionnels

La formulation des questions sous le format PICO est la première étape de la rédaction d'une RPC qui doit être réalisée avant que ne débutent les réunions du groupe. Ceci implique qu'une première recherche bibliographique ait été effectuée afin de déterminer s'il existe ou non de quoi évaluer dans la littérature la qualité de la preuve scientifique. L'absence de données sur certaines questions conduira le plus souvent à ne pas formuler les questions PICO en relation puisque l'on sait que l'on ne pourra pas y répondre. Étant donné qu'est abandonnée la présentation des RPC à la manière d'un ouvrage de référence, un nombre important de questions, qui était abordé précédemment car faisant partie de la logique éditoriale du chapitre, ne le sera plus. Ces questions, mal systématisées et trop ouvertes, aboutissaient fréquemment à la rédaction de recommandations sous la forme d'accord professionnel. La formulation des questions au format PICO réduira considérablement leur nombre. Toutefois, certaines questions, malgré l'absence de littérature sur le sujet, nécessiteront parfois, tout de même, un positionnement du CNGOF, et conduiront à la rédaction de recommandation sous la forme d'accord professionnel, correspondant au *Good Practice Point* ou *Statement* des Anglo-Saxons. Ceux-ci devront être limités au maximum, et ne pas dépasser si possible 20 % du nombre total de recommandations. Une RPC dont la majorité des recommandations relèverait d'accords professionnels, ne devrait pas porter le nom de RPC car la logique des accords d'experts du groupe de travail tel qu'il a été mis en place pour une démarche RPC ne constitue pas une démarche valide pour émettre des recommandations. Au contraire, lorsque l'analyse préliminaire de la littérature montre l'absence de données sur la plupart des questions c'est à une méthodologie de type « consensus formalisé d'experts » qu'il faut avoir recours.

II.c. Disparition des textes longs et des textes courts

La deuxième conséquence importante de la formulation des questions au format PICO est la disparition des textes dits « longs » publiés dans la revue de Gynécologie Obstétrique Fertilité & Sénologie. En effet, une RPC ne comprendra au maximum qu'une trentaine de questions PICO à laquelle répondront les experts. L'argumentaire exposant l'évaluation de la qualité de la preuve permettant de répondre

à cette question et conduisant à une recommandation tiendra le plus souvent en 1 ou 2 pages par question faisant disparaître la nécessité de publier des textes longs. Par conséquent, les textes courts, synthèse des textes longs, disparaîtront eux aussi.

II.d. Nouvelles modalités de publication des RPC

Les recommandations du CNGOF feront toujours l'objet d'une publication, en français et en anglais. Elles seront désormais présentées sous la forme d'un texte unique qui sera constitué par l'ensemble des argumentaires ayant permis de répondre aux questions PICO, accompagné des tableaux GRADE correspondants fournis en annexe au format électronique. Il sera aussi exposé dans ce texte le processus organisationnel ayant conduit et accompagné la rédaction de la RPC comme stipulé dans la grille AGREE II, et le lecteur devra être capable d'identifier le respect des 23 critères précédemment mentionnés. L'adhérence à ces critères de qualité permettra une meilleure reconnaissance internationale du travail du CNGOF.

II.e. Présentation aux Journées Paris Santé Femmes

Aux Journées Paris Santé Femmes qui remplaceront les journées du CNGOF à partir du mois de janvier 2020, les experts viendront présenter l'argumentaire des questions PICO auxquelles ils auront eu à répondre et les recommandations correspondantes. La durée des présentations en sera réduite et les échanges avec la salle devraient par conséquent être beaucoup plus nombreux et fructueux, à la manière de ce qui existait auparavant sous la forme des « rencontres avec les experts ».

II.f. Révision / réactualisation des RPC existantes

Les révisions de RPC existantes rédigées selon le système HAS se feront par l'introduction de questions PICO correspondant aux domaines de la RPC pour lesquels il existe des nouveautés importantes justifiant donc une réévaluation de la littérature scientifique. Si besoin, c'est la RPC dans son ensemble qui sera révisée, là encore uniquement sous la forme de questions PICO.

III. PLACE DES RPC DANS LA STRATÉGIE D'ACCOMPAGNEMENT DES PROFESSIONNELS POUR LA PRISE DE DÉCISIONS MÉDICALES

La rédaction d'une RPC fait souvent suite à la saisie du bureau du CA du CNGOF par des acteurs de santé ou des usagers, désirant un positionnement du collègue sur la prise en charge d'une pathologie afin de les aider dans leur prise de décisions. La réponse à toutes ces requêtes ne pouvait déjà pas toujours conduire à la rédaction de RPC lorsqu'était en vigueur la méthodologie HAS, le passage au système GRADE risque encore d'accentuer ce phénomène. En effet, lorsque la littérature est insuffisante, la réponse à une question PICO n'est pas possible. La réponse à la demande faite au CA peut alors prendre la forme d'une méthodologie de type « consensus formalisé d'experts » [19].

III.a. Le consensus formalisé d'experts

La méthode du consensus formalisé d'experts vise à dégager un consensus sur un sujet pour lequel la littérature scientifique est inexistante ou n'apporte pas de preuves de qualité suffisante permettant d'émettre des recommandations. Cette méthode est particulièrement adaptée pour répondre aux questions aboutissant dans l'ancien modèle des RPC à des accords professionnels. Pour couper définitivement avec la terminologie passée et pour que cette démarche ne soit pas assimilée à la démarche RPC, ces recommandations basées sur des accords entre professionnels experts du domaine seront désormais appelés conseils de bonne pratique (CBP) (*Good practice point* des Anglo-Saxons). Un consensus formalisé d'experts sera donc constitué d'une très grande majorité de conseils de bonne pratique sur un sujet donné pour lesquels il n'existe pas ou peu de littérature permettant d'établir une qualité de la preuve (par exemple la modalité de la sortie précoce après accouchement, la réhabilitation après chirurgie, ou bien l'indication de la préservation ovarienne pour pathologie bénigne, ou encore la prise en charge des complications des implants vaginaux). Par définition, elle ne pourra donc pas être redondante thématiquement avec une RPC. Les conseils de bonne pratique ne comporteront pas les termes : « il est recommandé de », « il est recommandé de ne pas », mais « il est proposé de », « il est proposé de ne pas ». L'ensemble des conseils de bonne pratique tiendra dans un argumentaire d'une vingtaine de page

destiné à être publié sous la forme d'un article en français et possiblement aussi en anglais.

Le CS décidera si une thématique proposée peut donner lieu à la tenue d'un consensus formalisé d'experts. Si c'est le cas, le CS désignera, comme pour une RPC, un coordonnateur, un méthodologiste, qui désigneront à leur tour les membres du comité d'organisation. Dans la logique de cette méthodologie dite qualitative, il est important de s'assurer de l'équilibre entre les différentes écoles de pensée, et de tenir compte de l'ensemble des disciplines et professions concernées. Le consensus formalisé d'experts répond aux mêmes exigences d'indépendance éditoriale qu'une RPC et les DPIs des membres du groupe de travail seront aussi relevées. Le coordonnateur sera le garant de l'impartialité du CS et veillera à ce que le rendu des conseils de bonne pratique soit un reflet équilibré des débats.

C'est ce groupe, correspondant au comité d'organisation, qui sera chargé d'élaborer des propositions initiales de conseils de bonne pratique qui seront soumises à un processus de validation renforcé. En effet, compte tenu de l'absence de preuve scientifique venant étayer les propositions de conseils de bonne pratique élaborées par le comité d'organisation, leur validation sera réalisée par un panel d'experts extérieur selon un processus beaucoup plus formalisé que pour les RPC classiques, la méthode DELPHI modifiée [20]. La méthode DELPHI modifiée à deux tours comporte la constitution d'un groupe d'experts externes multidisciplinaires (et indépendants du comité d'organisation), comprenant des professionnels de santé, gynécologues-obstétriciens ou non, impliqués dans la prise en charge de la pathologie faisant l'objet du consensus formalisé, et éventuellement des « patients experts ». La méthode comprend deux tours de cotation par les experts extérieurs avec possibilité de modifications et/ou d'ajout des propositions suivie d'une réunion physique de discussion et de finalisation de l'écriture des conseils de bonne pratique [19]. Cette technique DELPHI modifiée, dans laquelle les séries de questionnaires sont suivies d'une réunion en face à face des intervenants, permet d'améliorer le processus décisionnel complexe et de clarifier le langage utilisé pour écrire chaque proposition. Elle possède une très grande validité pour l'élaboration de recommandations en l'absence de preuve scientifique [20].

La méthode DELPHI modifiée à deux tours a été préférée aux autres méthodes d'obtention d'un consensus en l'absence de preuve scientifique, comme les méthodes RAND/UCLA [21], de groupe nominal [21] ou la conférence de consensus [21]. En effet, en comparaison avec ces autres méthodes, la méthode DELPHI modifiée

à deux tours présente l'intérêt d'une validation par un nombre important d'experts, pour un coût modique, tout en respectant l'anonymat [21]. Le succès de la méthode DELPHI repose sur la pertinence du choix des experts qui devront provenir d'horizons les plus variés possibles, en respectant l'interdisciplinarité, la parité homme/femme, public/privé, Paris/régions, universitaire/non universitaire. Les conseils de bonne pratique qui en seront issus feront, par définition, l'objet d'un fort consensus.

III.b. L'opinion d'experts

Les avis formulés par le CNGOF visant à aider les professionnels dans la prise de décision médicale mais ne répondant pas aux critères d'une RPC (questions PICO, méthode GRADE, minorité de recommandations sous la forme d'accords professionnels) ou à ceux d'un consensus formalisé d'experts (méthode DELPHI, conseils de bonne pratique) seront intitulés opinion d'experts.

CONCLUSION

Le CNGOF va opérer un virage important en changeant la méthodologie de ses RPC. En adoptant les critères du consortium AGREE II et la cotation de la preuve selon la méthode GRADE, il augmentera la qualité et la transparence de l'ensemble du processus, délivrera des messages plus ciblés et facilement assimilables, pour, nous l'espérons, faciliter mieux encore les prises de décisions des professionnels de notre spécialité, au bénéfice espéré de la santé des femmes et de leurs enfants.

La nouvelle méthodologie des RPC du CNGOF a bien entendu vocation à être évolutive, selon les modifications que proposeront dans les années à venir les consortiums AGREE II et GRADE.

Déclaration publique d'intérêt

L. SENTILHES a exercé des activités de consultant pour le laboratoire Ferring dans les trois dernières années. Les autres auteurs ne déclarent pas de liens d'intérêt.

Bibliographie

- [1] Recommandations pour la pratique clinique. Infections cervico-vaginales et grossesse. Disponibles à : <http://www.cngof.fr/pratiques-cliniques/recommandations-pour-la-pratique-clinique/aperçu?path=RPC%2BCOLLEGE%252F1997%252FINFECTIONS%2BCERVICO-VAGINALES%2BET%2BGROSSESSE%2B%25281997%2529.pdf&i=21977>
- [2] Goffinet F, Mercier F, Teyssier V, Pierre F, Dreyfus M, Mignon A, *et al.* Postpartum haemorrhage: recommendations for clinical practice by the CNGOF. *Gynecol Obstet Fertil* 2005;33:268-74.
- [3] Deneux-Tharaux C, Dupont C, Colin C, Rabilloud M, Touzet S, Lansac J, *et al.* Multifaceted intervention to decrease the rate of severe postpartum haemorrhage: the PITHAGORE6 cluster-randomised controlled trial. *BJOG* 2010;117:1278-87.
- [4] AGREE Next Steps Consortium (2009). The AGREE II Instrument [Electronic version]. Disponible à : <http://www.agreetrust.org>
- [5] Burda BU, Chambers AR, Johnson JC. Appraisal of guidelines developed by the World Health Organization. *Public Health* 2014;128:444-74.
- [6] Hirsch M, Begum MR, Paniz É, Barker C, Davis CJ, Duffy J. Diagnosis and management of endometriosis: a systematic review of international and national guidelines. *BJOG* 2018;125:556-64.
- [7] Dahlke JD, Mendez-Figueroa H, Maggio L, Hauspurg AK, Sperling JD, Chauhan SP, *et al.* Prevention and management of postpartum hemorrhage: a comparison of 4 national guidelines. *Am J Obstet Gynecol* 2015;213:76.e1-76.e10.
- [8] WHO Handbook for guideline development (2nd edition). Disponible à : https://www.who.int/publications/guidelines/handbook_2nd_ed.pdf?ua=1
- [9] <http://www.crown-initiative.org/journals/>
- [10] <http://www.comet-initiative.org>
- [11] Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, *et al.* Grading quality of evidence and strength of recommendations. *BMJ* 2004;328:1490.
- [12] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter RY, Alonso-Coello P, *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6.
- [13] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist GE, Brozek J, *et al.* GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383-94.
- [14] Schmitz T, Kayem G, Senat MV, Sentilhes L. Rupture prématurée de membranes avant terme: Recommandations pour la pratique clinique - Méthodologie. *Gynecol Obstet Fertil Senol* 2018;46:996-7.
- [15] Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ* 2003;327:1459-61.
- [16] Brubaker L, Nygaard I, Richter HE, Visco A, Weber AM, Cundiff GW, Fine P, Ghetti C, Brown MB. Two-year outcomes after sacrocolpopexy with and without burch to prevent stress urinary incontinence. *Obstet Gynecol* 2008;112:49-55.
- [17] Neumann I, Santesso N, Akl EA, Rind DM, Vandvik PO, Alonso-Coello P, *et al.* A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. *J Clin Epidemiol* 2016;72:45-55.
- [18] Agoritsas T, Merglen A, Heen AF, Kristiansen A, Neumann I, Brito JP, *et al.* UpToDate adherence to GRADE criteria for strong recommendations: an analytical survey. *BMJ Open* 2017;7:e018593.
- [19] Bourree F, Michel P, Salmi LR. Consensus methods: review of original methods and their main alternatives used in public health. *Rev Epidemiol Sante Publique*. 2008;56:415-23.
- [20] Boukedi R, Sibony O, Goffinet F, Fauconnier A, Branger B, Alberti C. Quality indicators for continuous monitoring to improve maternal and infant health in maternity departments: a modified Delphi survey of an international multidisciplinary panel. *PLoS One*. 2013;8:e60663.
- [21] Humphrey-Murto S, Varpio L, Wood TJ, Gonsalves C, Ufholz LA, Mascioli K, Wang C, Foth T. The Use of the Delphi and Other Consensus Group Methods in Medical Education Research: A Review. *Acad Med* 2017;92:1491-8.